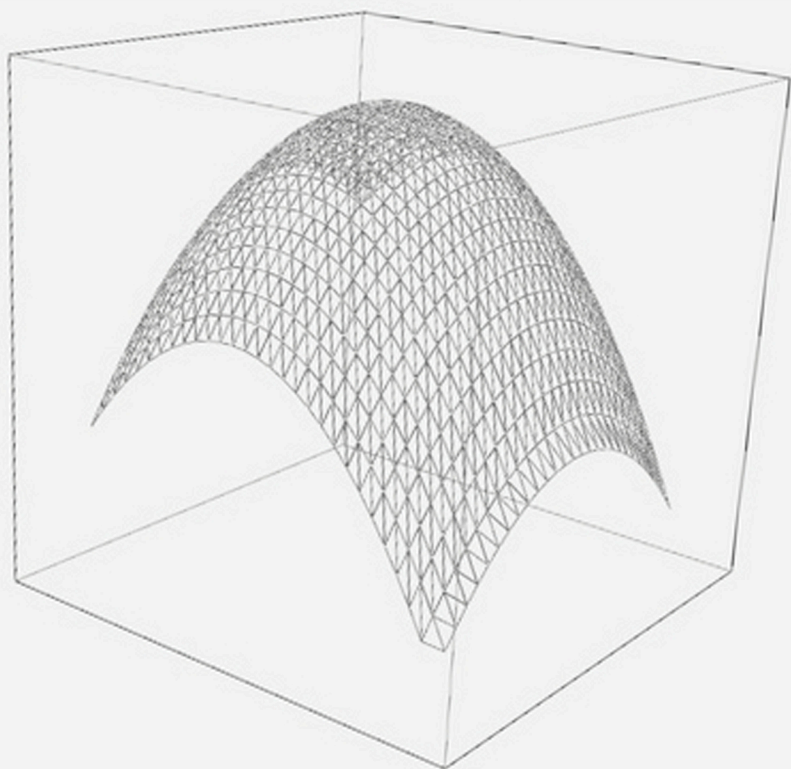


Wiley Series in Probability and Statistics

# RESPONSE SURFACE METHODOLOGY

PROCESS AND PRODUCT OPTIMIZATION  
USING DESIGNED EXPERIMENTS

FOURTH EDITION



RAYMOND H. MYERS  
DOUGLAS C. MONTGOMERY  
CHRISTINE M. ANDERSON-COOK

WILEY



# **RESPONSE SURFACE METHODOLOGY**

WILEY SERIES IN PROBABILITY AND STATISTICS

Established by WALTER A. SHEWHART and SAMUEL S. WILKS

Editors: *David J. Balding, Noel A. C. Cressie, Garrett M. Fitzmaurice, Geof H. Givens, Harvey Goldstein, Geert Molenberghs, David W. Scott, Adrian F. M. Smith, Ruey S. Tsay, Sanford Weisberg*

Editors Emeriti: *J. Stuart Hunter, Iain M. Johnstone, Joseph B. Kadane, Jozef L. Teugels*

A complete list of the titles in this series appears at the end of this volume.

# **RESPONSE SURFACE METHODOLOGY**

---

## **Process and Product Optimization Using Designed Experiments**

Fourth Edition

**RAYMOND H. MYERS**

Virginia Polytechnic University, Department of Statistics, Blacksburg, VA

**DOUGLAS C. MONTGOMERY**

Arizona State University, Department of Industrial Engineering, Tempe, AZ

**CHRISTINE M. ANDERSON-COOK**

Los Alamos National Laboratory, Los Alamos, NM

**WILEY**

Copyright © 2016 by John Wiley & Sons, Inc. All rights reserved

Published by John Wiley & Sons, Inc., Hoboken, New Jersey  
Published simultaneously in Canada

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at [www.copyright.com](http://www.copyright.com). Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permission>.

**Limit of Liability/Disclaimer of Warranty:** While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic format. For more information about Wiley products, visit our web site at [www.wiley.com](http://www.wiley.com).

***Library of Congress Cataloging-in-Publication Data:***

Myers, Raymond H.

Response surface methodology : process and product optimization using designed experiments.

– 4th ed. / Raymond H. Myers, Douglas C. Montgomery, Christine M. Anderson-Cook.

p. cm. – (Wiley series in probability and statistics)

Includes bibliographical references and index.

ISBN 978-1-118-91601-8 (cloth)

1. Experimental design. 2. Response surfaces (Statistics). I. Montgomery, Douglas C.

II. Anderson-Cook, Christine M. III. Title.

QA279.M94 2016

519.507–dc22

2016019012

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

# CONTENTS

<b>Preface</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Response Surface Methodology, 1	
1.1.1 Approximating Response Functions, 2	
1.1.2 The Sequential Nature of RSM, 7	
1.1.3 Objectives and Typical Applications of RSM, 9	
1.1.4 RSM and the Philosophy of Quality Improvement, 11	
1.2 Product Design and Formulation (Mixture Problems), 11	
1.3 Robust Design and Process Robustness Studies, 12	
1.4 Useful References on RSM, 12	
<b>2 Building Empirical Models</b>	<b>13</b>
2.1 Linear Regression Models, 13	
2.2 Estimation of the Parameters in Linear Regression Models, 14	
2.3 Properties of the Least Squares Estimators and Estimation of $\sigma^2$ , 22	
2.4 Hypothesis Testing in Multiple Regression, 24	
2.4.1 Test for Significance of Regression, 24	
2.4.2 Tests on Individual Regression Coefficients and Groups of Coefficients, 27	
2.5 Confidence Intervals in Multiple Regression, 31	
2.5.1 Confidence Intervals on the Individual Regression Coefficients $\beta$ , 32	
2.5.2 A Joint Confidence Region on the Regression Coefficients $\beta$ , 32	
2.5.3 Confidence Interval on the Mean Response, 33	
2.6 Prediction of New Response Observations, 35	

2.7	Model Adequacy Checking, 36	
2.7.1	Residual Analysis, 36	
2.7.2	Scaling Residuals, 38	
2.7.3	Influence Diagnostics, 42	
2.7.4	Testing for Lack of Fit, 43	
2.8	Fitting a Second-Order Model, 47	
2.9	Qualitative Regressor Variables, 55	
2.10	Transformation of the Response Variable, 61	
	Exercises, 66	
<b>3</b>	<b>Two-Level Factorial Designs</b>	<b>81</b>
3.1	Introduction, 81	
3.2	The $2^2$ Design, 82	
3.3	The $2^3$ Design, 94	
3.4	The General $2^k$ Design, 103	
3.5	A Single Replicate of the $2^k$ Design, 108	
3.6	$2^k$ Designs are Optimal Designs, 125	
3.7	The Addition of Center Points to the $2^k$ Design, 130	
3.8	Blocking in the $2^k$ Factorial Design, 135	
3.8.1	Blocking in the Replicated Design, 135	
3.8.2	Confounding in the $2^k$ Design, 137	
3.9	Split-Plot Designs, 141	
	Exercises, 146	
<b>4</b>	<b>Two-Level Fractional Factorial Designs</b>	<b>161</b>
4.1	Introduction, 161	
4.2	The One-Half Fraction of the $2^k$ Design, 162	
4.3	The One-Quarter Fraction of the $2^k$ Design, 174	
4.4	The General $2^{k-p}$ Fractional Factorial Design, 184	
4.5	Resolution III Designs, 188	
4.6	Resolution IV and V Designs, 197	
4.7	Alias Structures in Fractional Factorial and Other Designs, 198	
4.8	Nonregular Fractional Factorial Designs, 200	
4.8.1	Nonregular Fractional Factorial Designs for 6, 7, and 8 Factors in 16 Runs, 203	
4.8.2	Nonregular Fractional Factorial Designs for 9 Through 14 Factors in 16 Runs, 209	
4.8.3	Analysis of Nonregular Fractional Factorial Designs, 213	
4.9	Fractional Factorial Split-Plot Designs, 216	
4.10	Summary, 219	
	Exercises, 220	
<b>5</b>	<b>Process Improvement with Steepest Ascent</b>	<b>233</b>
5.1	Determining the Path of Steepest Ascent, 234	
5.1.1	Development of the Procedure, 234	
5.1.2	Practical Application of the Method of Steepest Ascent, 237	



5.2	Consideration of Interaction and Curvature, 241	
5.2.1	What About a Second Phase?, 244	
5.2.2	What Happens Following Steepest Ascent?, 244	
5.3	Effect of Scale (Choosing Range of Factors), 245	
5.4	Confidence Region for Direction of Steepest Ascent, 247	
5.5	Steepest Ascent Subject to a Linear Constraint, 250	
5.6	Steepest Ascent in a Split-Plot Experiment, 254	
	Exercises, 262	
<b>6</b>	<b>The Analysis of Second-Order Response Surfaces</b>	<b>273</b>
6.1	Second-Order Response Surface, 273	
6.2	Second-Order Approximating Function, 274	
6.2.1	The Nature of the Second-Order Function and Second-Order Surface, 274	
6.2.2	Illustration of Second-Order Response Surfaces, 276	
6.3	A Formal Analytical Approach to the Second-Order Model, 277	
6.3.1	Location of the Stationary Point, 278	
6.3.2	Nature of the Stationary Point (Canonical Analysis), 278	
6.3.3	Ridge Systems, 282	
6.3.4	Role of Contour Plots, 286	
6.4	Ridge Analysis of the Response Surface, 289	
6.4.1	Benefits of Ridge Analysis, 290	
6.4.2	Mathematical Development of Ridge Analysis, 291	
6.5	Sampling Properties of Response Surface Results, 296	
6.5.1	Standard Error of Predicted Response, 296	
6.5.2	Confidence Region on the Location of the Stationary Point, 299	
6.5.3	Use and Computation of the Confidence Region on the Location of the Stationary Point, 300	
6.5.4	Confidence Intervals on Eigenvalues in Canonical Analysis, 304	
6.6	Further Comments Concerning Response Surface Analysis, 307	
	Exercises, 307	
<b>7</b>	<b>Multiple Response Optimization</b>	<b>325</b>
7.1	Balancing Multiple Objectives, 325	
7.2	Strategies for Multiple Response Optimization, 338	
7.2.1	Overlaying Contour Plots, 339	
7.2.2	Constrained Optimization, 340	
7.2.3	Desirability Functions, 341	
7.2.4	Pareto Front Optimization, 343	
7.2.5	Other Options for Optimization, 349	
7.3	A Sequential Process for Optimization—DMRCS, 350	
7.4	Incorporating Uncertainty of Response Predictions into Optimization, 352	
	Exercises, 357	

<b>8</b>	<b>Design of Experiments for Fitting Response Surfaces—I</b>	<b>369</b>
8.1	Desirable Properties of Response Surface Designs, 369	
8.2	Operability Region, Region of Interest, and Metrics for Desirable Properties, 371	
8.2.1	Metrics for Desirable Properties, 372	
8.2.2	Model Inadequacy and Model Bias, 373	
8.3	Design of Experiments for First-Order Models and First-Order Models with Interactions, 375	
8.3.1	The First-Order Orthogonal Design, 376	
8.3.2	Orthogonal Designs for Models Containing Interaction, 378	
8.3.3	Other First-Order Orthogonal Designs—The Simplex Design, 381	
8.3.4	Definitive Screening Designs, 385	
8.3.5	Another Variance Property—Prediction Variance, 389	
8.4	Designs for Fitting Second-Order Models, 393	
8.4.1	The Class of Central Composite Designs, 393	
8.4.2	Design Moments and Property of Rotatability, 399	
8.4.3	Rotatability and the CCD, 403	
8.4.4	More on Prediction Variance—Scaled, Unscaled, and Estimated, 406	
8.4.5	The Face-Centered Cube in Cuboidal Regions, 408	
8.4.6	Choosing between Spherical and Cuboidal Regions, 411	
8.4.7	The Box–Behnken Design, 413	
8.4.8	Definitive Screening Designs for Fitting Second-Order Models, 417	
8.4.9	Orthogonal Blocking in Second-Order Designs, 422	
	Exercises, 434	
<b>9</b>	<b>Experimental Designs for Fitting Response Surfaces—II</b>	<b>451</b>
9.1	Designs that Require a Relatively Small Run Size, 452	
9.1.1	The Hoke Designs, 452	
9.1.2	Koshal Design, 454	
9.1.3	Hybrid Designs, 455	
9.1.4	The Small Composite Design, 458	
9.1.5	Some Saturated or Near-Saturated Cuboidal Designs, 462	
9.1.6	Equiradial Designs, 463	
9.2	General Criteria for Constructing, Evaluating, and Comparing Designed Experiments, 465	
9.2.1	Practical Design Optimality, 467	
9.2.2	Use of Design Efficiencies for Comparison of Standard Second-Order Designs, 474	
9.2.3	Graphical Procedure for Evaluating the Prediction Capability of an RSM Design, 477	
9.3	Computer-Generated Designs in RSM, 488	
9.3.1	Important Relationship Between Prediction Variance and Design Augmentation for <i>D</i> -Optimality, 491	
9.3.2	Algorithms for Computer-Generated Designs, 494	
9.3.3	Comparison of <i>D</i> -, <i>G</i> -, and <i>I</i> -Optimal Designs, 497	

- 9.3.4 Illustrations Involving Computer-Generated Design, 499
  - 9.3.5 Computer-Generated Designs Involving Qualitative Variables, 508
  - 9.4 Multiple Objective Computer-Generated Designs for RSM, 517
    - 9.4.1 Pareto Front Optimization for Selecting a Design, 518
    - 9.4.2 Pareto Aggregating Point Exchange Algorithm, 519
    - 9.4.3 Using DMRCs for Design Optimization, 520
  - 9.5 Some Final Comments Concerning Design Optimality and Computer-Generated Design, 525
- Exercises, 527

## **10 Advanced Topics in Response Surface Methodology 543**

- 10.1 Effects of Model Bias on the Fitted Model and Design, 543
  - 10.2 A Design Criterion Involving Bias and Variance, 547
    - 10.2.1 The Case of a First-Order Fitted Model and Cuboidal Region, 550
    - 10.2.2 Minimum Bias Designs for a Spherical Region of Interest, 556
    - 10.2.3 Simultaneous Consideration of Bias and Variance, 558
    - 10.2.4 How Important Is Bias?, 558
  - 10.3 Errors in Control of Design Levels, 560
  - 10.4 Experiments with Computer Models, 563
    - 10.4.1 Design for Computer Experiments, 567
    - 10.4.2 Analysis for Computer Experiments, 570
    - 10.4.3 Combining Information from Physical and Computer Experiments, 574
  - 10.5 Minimum Bias Estimation of Response Surface Models, 575
  - 10.6 Neural Networks, 579
  - 10.7 Split-Plot Designs for Second-Order Models, 581
  - 10.8 RSM for Non-Normal Responses—Generalized Linear Models, 591
    - 10.8.1 Model Framework: The Link Function, 592
    - 10.8.2 The Canonical Link Function, 593
    - 10.8.3 Estimation of Model Coefficients, 593
    - 10.8.4 Properties of Model Coefficients, 595
    - 10.8.5 Model Deviance, 595
    - 10.8.6 Overdispersion, 597
    - 10.8.7 Examples, 598
    - 10.8.8 Diagnostic Plots and Other Aspects of the GLM, 605
- Exercises, 609

## **11 Robust Parameter Design and Process Robustness Studies 619**

- 11.1 Introduction, 619
- 11.2 What is Parameter Design?, 619
  - 11.2.1 Examples of Noise Variables, 620
  - 11.2.2 An Example of Robust Product Design, 621
- 11.3 The Taguchi Approach, 622
  - 11.3.1 Crossed Array Designs and Signal-to-Noise Ratios, 622
  - 11.3.2 Analysis Methods, 625
  - 11.3.3 Further Comments, 630

11.4	The Response Surface Approach, 631	
11.4.1	The Role of the Control $\times$ Noise Interaction, 631	
11.4.2	A Model Containing Both Control and Noise Variables, 635	
11.4.3	Generalization of Mean and Variance Modeling, 638	
11.4.4	Analysis Procedures Associated with the Two Response Surfaces, 642	
11.4.5	Estimation of the Process Variance, 651	
11.4.6	Direct Variance Modeling, 655	
11.4.7	Use of Generalized Linear Models, 657	
11.5	Experimental Designs For RPD and Process Robustness Studies, 661	
11.5.1	Combined Array Designs, 661	
11.5.2	Second-Order Designs, 663	
11.5.3	Other Aspects of Design, 665	
11.6	Dispersion Effects in Highly Fractionated Designs, 672	
11.6.1	The Use of Residuals, 673	
11.6.2	Further Diagnostic Information from Residuals, 674	
11.6.3	Further Comments Concerning Variance Modeling, 680	
	Exercises, 684	

**12 Experiments with Mixtures 693**

12.1	Introduction, 693
12.2	Simplex Designs and Canonical Mixture Polynomials, 696
12.2.1	Simplex Lattice Designs, 696
12.2.2	The Simplex-Centroid Design and Its Associated Polynomial, 704
12.2.3	Augmentation of Simplex Designs with Axial Runs, 707
12.3	Response Trace Plots, 716
12.4	Reparameterizing Canonical Mixture Models to Contain A Constant Term ( $\beta_0$ ), 716
	Exercises, 720

**13 Other Mixture Design and Analysis Techniques 731**

13.1	Constraints on the Component Proportions, 731
13.1.1	Lower-Bound Constraints on the Component Proportions, 732
13.1.2	Upper-Bound Constraints on the Component Proportions, 743
13.1.3	Active Upper- and Lower-Bound Constraints, 747
13.1.4	Multicomponent Constraints, 758
13.2	Mixture Experiments Using Ratios of Components, 759
13.3	Process Variables in Mixture Experiments, 763
13.3.1	Mixture-Process Model and Design Basics, 763
13.3.2	Split-Plot Designs for Mixture-Process Experiments, 767
13.3.3	Robust Parameter Designs for Mixture-Process Experiments, 778
13.4	Screening Mixture Components, 783
	Exercises, 785

<b>Appendix 1</b>	<b>Moment Matrix of a Rotatable Design</b>	<b>797</b>
<b>Appendix 2</b>	<b>Rotatability of a Second-Order Equiradial Design</b>	<b>803</b>
<b>References</b>		<b>807</b>
<b>Index</b>		<b>821</b>



# PREFACE

This book deals with the exploration and optimization of response surfaces. This is a problem faced by experimenters in many technical fields, where, in general, the response variable of interest is  $y$  and there is a set of predictor variables  $x_1, x_2, \dots, x_k$ . For example,  $y$  might be the viscosity of a polymer and  $x_1, x_2$ , and  $x_3$  might be the process variables reaction time, reactor temperature, and the catalyst feed rate in the process. In some systems the nature of the relationship between  $y$  and the  $x$ 's might be known "exactly," based on the underlying engineering, chemical, or physical principles. Then we could write a model of the form  $y = g(x_1, x_2, \dots, x_k) + \epsilon$ , where  $g$  is a known (often nonlinear) function and  $\epsilon$  represents the "error" in the system. This type of relationship is often called a mechanistic model. We consider the more common situation where the underlying mechanism is not fully understood, and the experimenter must approximate the unknown function  $g$  with an appropriate empirical model  $y = f(x_1, x_2, \dots, x_k) + \epsilon$ . Usually the function  $f$  is a low-order polynomial, typically either a first-order or second-order polynomial. This empirical model is called a response surface model.

Identifying and fitting an appropriate response surface model from experimental data requires some knowledge of statistical experimental design fundamentals, regression modeling techniques, and elementary optimization methods. This book integrates all three of these topics into what has been popularly called response surface methodology (RSM). We assume that the reader has some previous exposure to statistical methods and matrix algebra. Formal coursework in basic principles of experimental design and regression analysis would be helpful, but are not essential, because the important elements of these topics are presented early in the text. We have used this book in a graduate-level course on RSM for statisticians, engineers, and chemical/physical scientists. We have also used it in industrial short courses and seminars for individuals with a wide variety of technical backgrounds.

This fourth edition is a substantial revision of the book. We have rewritten many sections to incorporate new topics and material, ideas, and examples and to more fully explain some topics that were only briefly mentioned in previous editions and added a new chapter. We have also continued to integrate the computer tightly into the presentation, relying on JMP

and Design-Expert for much of the computing, but also continuing to employ SAS for a few applications.

Chapters 1 through 4 contain the preliminary material essential to studying RSM. Chapter 1 is an introduction to the general field of RSM, describing typical applications such as (a) finding the levels of process variables that optimize a response of interest or (b) discovering what levels of these process variables result in a product satisfying certain requirements or specifications on responses such as yield, molecular weight, purity, or viscosity. Chapter 2 is a summary of regression methods useful in response surface work, focusing on the basic ideas of least squares model fitting, diagnostic checking, and inference for the linear regression model. Chapters 3 and 4 describe two-level factorial and fractional factorial designs. These designs are essential for factor screening or identifying the correct set of process variables to use in the RSM study. They are also basic building blocks for many of the response surface designs discussed later in the text. New topics in these chapters include additional information on analysis of unreplicated design and the use of nonregular fractional factorial designs in factor screening experiments.

Chapter 5 presents the method of steepest ascent, a simple but powerful optimization procedure used at the early stages of RSM to move the process from a region of relatively poor performance to one of greater potential. Chapter 6 introduces the analysis and optimization of a second-order response surface model. Both graphical and numerical techniques are presented. A new chapter, Chapter 7, describes techniques for the simultaneous optimization of several responses, a common problem in the application of RSM when experimenters want to extract information about several different characteristics of the process or product.

Chapters 8 and 9 present detailed information on the choice of experimental designs for fitting response surface models. Chapter 8 considers suitable metrics for evaluating the quality of estimation and prediction of a design and focuses on standard designs, including the central composite and Box–Behnken designs, and the important topic of blocking a response surface design. A new topic in this chapter is the important new class of designs called Definitive Screening Designs. Chapter 9 covers small response surface designs, design optimality criteria, and methods for evaluation of the prediction properties of response surface models constructed from various designs. We focus on variance dispersion graphs and fraction of design space plots, which are very important ways to summarize prediction properties. The section on computer generated designs has been expanded to describe some of the common algorithms used to generate designs as well as how to incorporate qualitative factors and multiple design criteria.

Chapter 10 contains more advanced RSM topics, including the use of mean square error as a design criterion, the effect of errors in controllable variables, split-plot type designs in a response surface setting, and the use of generalized linear models in the analysis of response surface experiments. The section on design and analysis of experiments for computer models has been updated to reflect new developments in this growing area. Chapter 11 describes how the problem of robust parameter design originally proposed by Taguchi can be efficiently solved in the RSM framework. We show how RSM not only makes the original problem posed by Taguchi easier to solve, but also provides much more information to the analyst about process or system performance. This chapter also contains much information on robust parameter design and process robustness studies. Chapters 12 and 13 present techniques for designing and analyzing experiments that involve mixtures. A mixture experiment is a special type of response surface experiment in which the design factors are the components or ingredients of a mixture, and the response depends on the



proportions of the ingredients that are present. Some new design techniques are introduced in these chapters.

We have provided expanded end-of-chapter problems and an updated reference section. The previous three editions of the text were written to emphasize methods that are useful in industry and that we have found useful in our own consulting experience. We have continued that applied focus in this new edition, though much new material has been added. We develop enough of the underlying theory to allow the reader to gain an understanding of the assumptions and conditions necessary to successfully apply RSM.

We are grateful to many individuals that have contributed meaningfully to this book. In particular, Dr. Bradley Jones, Mr. Pat Whitcomb, Dr. Geoff Vining, Dr. Soren Bisgaard, Dr. Connie Borrer, Dr. Rachel Silivstrini, Dr. Scott Kowalski, Dr. Dennis Lin, Dr. George Runger, Dr. Enrique Del Castillo, Dr. Lu Lu and Dr. Jessica Chapman made many useful suggestions. Dr. Matt Carlyle and Dr. Enrique Del Castillo also provided some figures that were most helpful. We also thank the many classes of graduate students that have studied from the book, as well as the instructors that have used the book. They have made many helpful comments and suggestions to improve the clarity of the presentation. We have tried to incorporate many of their suggestions. We also thank John Wiley & Sons for permission to use and adapt copyrighted material.

*Blacksburg, Virginia*  
*Tempe, Arizona*  
*Los Alamos, New Mexico*

RAYMOND H. MYERS  
DOUGLAS C. MONTGOMERY  
CHRISTINE M. ANDERSON-COOK



---

# 1

---

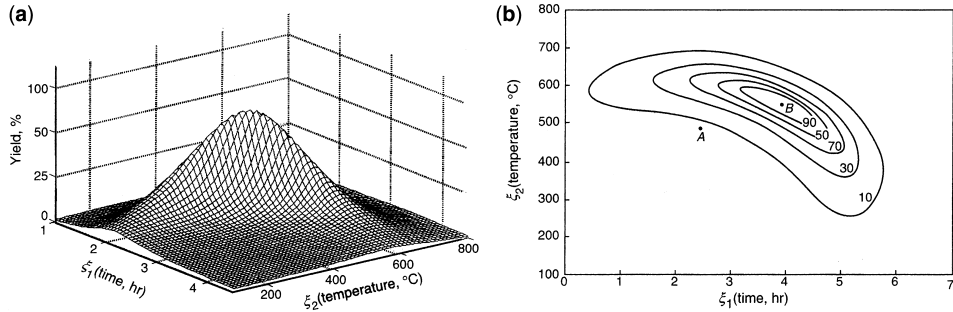
## INTRODUCTION

### 1.1 RESPONSE SURFACE METHODOLOGY

Response surface methodology (RSM) is a collection of statistical and mathematical techniques useful for developing, improving, and optimizing processes. It also has important applications in the design, development, and formulation of new products, as well as in the improvement of existing product designs.

The most extensive applications of RSM are in the industrial world, particularly in situations where several input variables potentially influence performance measures or quality characteristics of the product or process. These performance measures or quality characteristics are called the **response**. They are typically measured on a continuous scale, although attribute responses, ranks, and sensory responses are not unusual. Most real-world applications of RSM will involve more than one response. The input variables are sometimes called **independent variables**, and they are subject to the control of the engineer or scientist, at least for purposes of a test or an experiment.

Figure 1.1 shows graphically the relationship between the response variable yield ( $y$ ) in a chemical process and the two process variables (or independent variables) reaction time ( $\xi_1$ ) and reaction temperature ( $\xi_2$ ). Note that for each value of  $\xi_1$  and  $\xi_2$  there is a corresponding value of yield  $y$  and that we may view these values of the response yield as a surface lying above the time–temperature plane, as in Fig. 1.1a. It is this graphical perspective of the problem environment that has led to the term **response surface methodology**. It is also convenient to view the response surface in the two-dimensional time–temperature plane, as in Fig. 1.1b. In this presentation we are looking down at the time–temperature plane and connecting all points that have the same yield to produce contour lines of constant response. This type of display is called a **contour plot**.



**Figure 1.1** (a) A theoretical response surface showing the relationship between yield of a chemical process and the process variables reaction time ( $\xi_1$ ) and reaction temperature ( $\xi_2$ ). (b) A contour plot of the theoretical response surface.

Clearly, if we could easily construct the graphical displays in Fig. 1.1, optimization of this process would be very straightforward. By inspection of the plot, we note that yield is maximized in the vicinity of time  $\xi_1 = 4$  hr and temperature  $\xi_2 = 525^\circ\text{C}$ . Unfortunately, in most practical situations, the true response function in Fig. 1.1 is unknown. The field of response surface methodology consists of the experimental strategies for exploring the space of the process or independent variables (here the variables  $\xi_1$  and  $\xi_2$ ), empirical statistical modeling to develop an appropriate approximating relationship between the yield and the process variables, and optimization methods for finding the levels or values of the process variables  $\xi_1$  and  $\xi_2$  that produce desirable values of the responses (in this case that maximize yield).

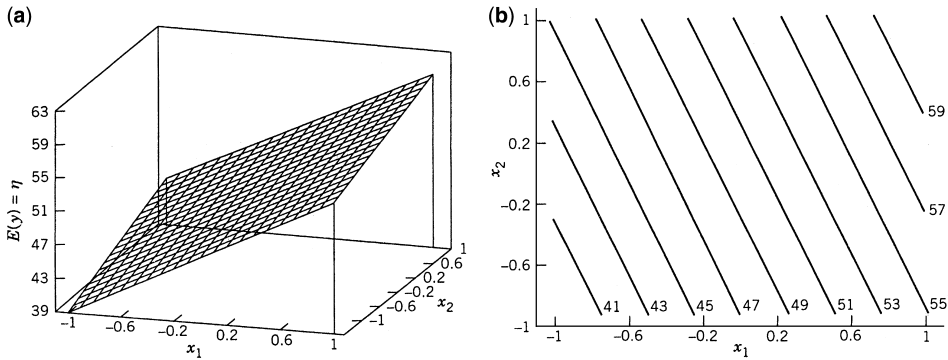
### 1.1.1 Approximating Response Functions

In general, suppose that the scientist or engineer (whom we will refer to as the **experimenter**) is concerned with a product, process, or system involving a response  $y$  that depends on the controllable input variables  $\xi_1, \xi_2, \dots, \xi_k$ . These input variables are also sometimes called factors, **independent variables**, or process variables. The actual relationship can be written

$$y = f(\xi_1, \xi_2, \dots, \xi_k) + \varepsilon \quad (1.1)$$

where the form of the true response function  $f$  is unknown and perhaps very complicated, and  $\varepsilon$  is a term that represents other sources of variability not accounted for in  $f$ . Thus  $\varepsilon$  includes effects such as measurement error on the response, other sources of variation that are inherent in the process or system (background noise, or common/special cause variation in the language of statistical process control), the effect of other (possibly unknown) variables, and so on. We will treat  $\varepsilon$  as a **statistical error**, often assuming it to have a normal distribution with mean zero and variance  $\sigma^2$ . If the mean of  $\varepsilon$  is zero, then

$$\begin{aligned} E(y) &\equiv \eta = E[f(\xi_1, \xi_2, \dots, \xi_k)] + E(\varepsilon) \\ &= f(\xi_1, \xi_2, \dots, \xi_k) \end{aligned} \quad (1.2)$$



**Figure 1.2** (a) Response surface for the first-order model  $\eta = 50 + 8x_1 + 3x_2$ . (b) Contour plot for the first-order model.

The variables  $\xi_1, \xi_2, \dots, \xi_k$  in Equation 1.2 are usually called the **natural variables**, because they are expressed in the natural units of measurement, such as degrees Celsius ( $^{\circ}\text{C}$ ), pounds per square inch (psi), or grams per liter for concentration. In much RSM work it is convenient to transform the natural variables to **coded variables**  $x_1, x_2, \dots, x_k$ , which are usually defined to be dimensionless with mean zero and the same spread or standard deviation. In terms of the coded variables, the true response function (1.2) is now written as

$$\eta = f(x_1, x_2, \dots, x_k) \quad (1.3)$$

Because the form of the true response function  $f$  is unknown, we must approximate it. In fact, successful use of RSM is critically dependent upon the experimenter's ability to develop a suitable approximation for  $f$ . Usually, a **low-order polynomial** in some relatively small region of the independent variable space is appropriate. In many cases, either a **first-order** or a **second-order** model is used. For the case of two independent variables, the first-order model in terms of the coded variables is

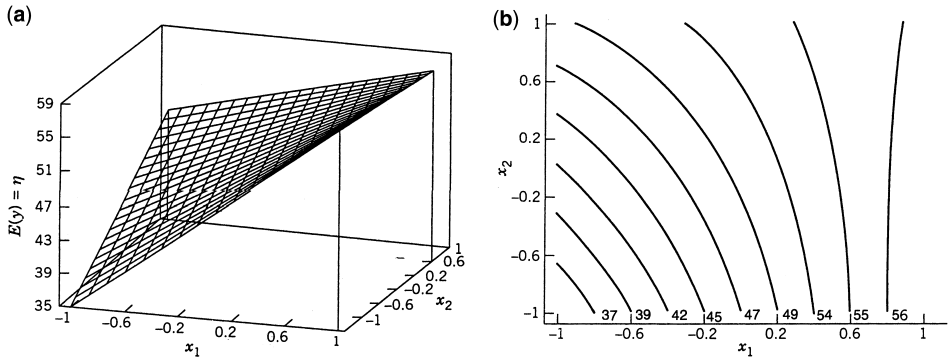
$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \quad (1.4)$$

Figure 1.2 shows the three-dimensional response surface and the two-dimensional contour plot for a particular case of the first-order model, namely,

$$\eta = 50 + 8x_1 + 3x_2$$

In three dimensions, the response surface for  $y$  is a plane lying above the  $x_1, x_2$  space. The contour plot shows that the first-order model can be represented as parallel straight lines of constant response in the  $x_1, x_2$  plane.

The first-order model is likely to be appropriate when the experimenter is interested in approximating the true response surface over a relatively small region of the independent variable space in a location where there is little curvature in  $f$ . For example, consider a small region around the point A in Fig. 1.1b; the first-order model would likely be appropriate here.



**Figure 1.3** (a) Response surface for the first-order model with interaction  $\eta = 50 + 8x_1 + 3x_2 - 4x_1x_2$ . (b) Contour plot for the first-order model with interaction.

The form of the first-order model in Equation 1.4 is sometimes called a **main effects model**, because it includes only the main effects of the two variables  $x_1$  and  $x_2$ . If there is an **interaction** between these variables, it can be added to the model easily as follows:

$$\eta = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{12}x_1x_2 \quad (1.5)$$

This is the **first-order model with interaction**. Figure 1.3 shows the three-dimensional response surface and the contour plot for the special case

$$\eta = 50 + 8x_1 + 3x_2 - 4x_1x_2$$

Notice that adding the interaction term  $-4x_1x_2$  introduces curvature into the response function. This leads to different rates of change of the response as  $x_1$  is changed for different fixed values of  $x_2$ . Similarly, the rate of change in  $y$  across  $x_2$  varies for different fixed values of  $x_1$ .

Often the curvature in the true response surface is strong enough that the first-order model (even with the interaction term included) is inadequate. A **second-order model** will likely be required in these situations. For the case of two variables, the second-order model is

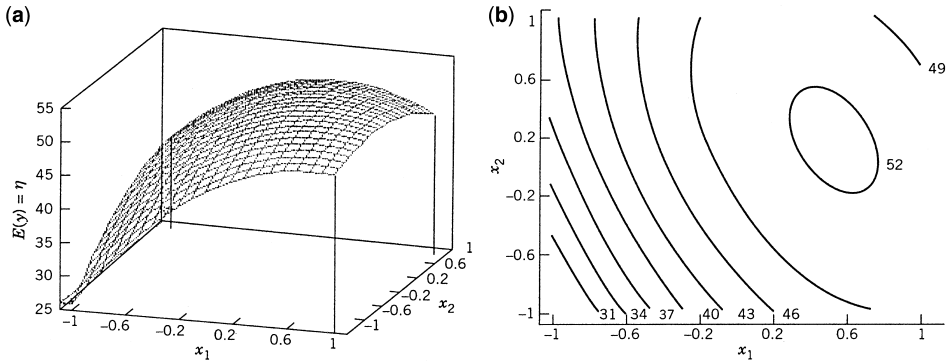
$$\eta = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{11}x_1^2 + \beta_{22}x_2^2 + \beta_{12}x_1x_2 \quad (1.6)$$

This model would likely be useful as an approximation to the true response surface in a relatively small region around the point  $B$  in Fig. 1.1b, where there is substantial curvature in the true response function  $f$ .

Figure 1.4 presents the response surface and contour plot for the special case of the second-order model

$$\eta = 50 + 8x_1 + 3x_2 - 7x_{11}^2 - 3x_{22}^2 - 4x_1x_2$$

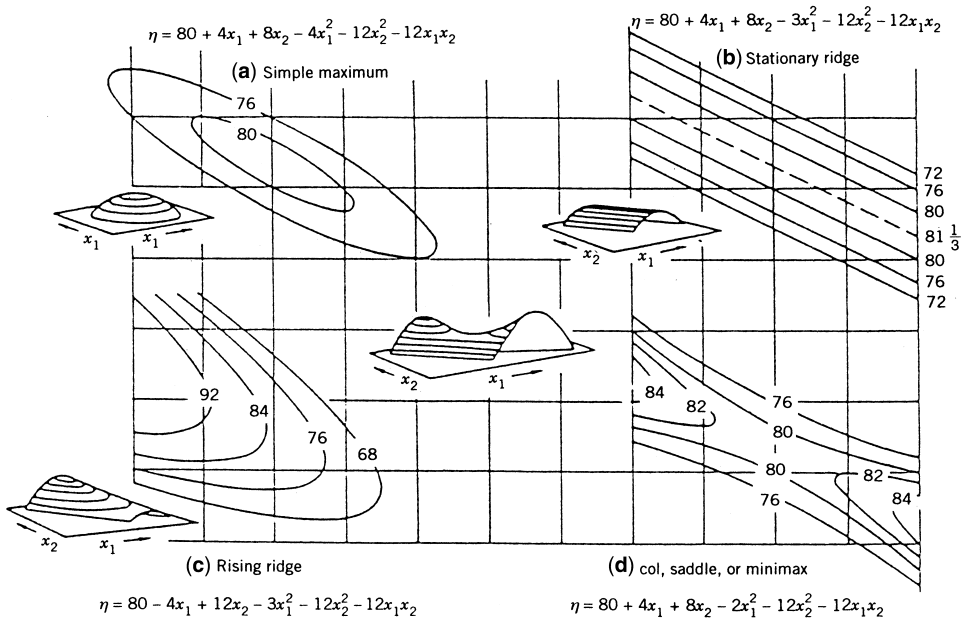
Notice the mound-shaped response surface and elliptical contours generated by this model. Such a response surface could arise in approximating a response such as yield, where we would expect to be operating near a maximum point on the surface.



**Figure 1.4** (a) Response surface for the second-order model  $\eta = 50 + 8x_1 + 3x_2 - 7x_1^2 - 3x_2^2 - 4x_1x_2$ . (b) Contour plot for the second-order model.

The second-order model is widely used in response surface methodology for several reasons. Among these are the following:

1. The second-order model is very **flexible**. It can take on a wide variety of functional forms, so it will often work well as an approximation to the true response surface. Figure 1.5 shows several different response surfaces and contour plots that can be generated by a second-order model.



**Figure 1.5** Some examples of types of surfaces defined by the second-order model in two variables  $x_1$  and  $x_2$ . (Adapted with permission from *Empirical Model Building and Response Surfaces*, G. E. P. Box and N. R. Draper, John Wiley & Sons, New York, 1987.)

2. It is **easy to estimate the parameters** (the  $\beta$ 's) in the second-order model. The method of least squares, which is presented in Chapter 2, can be used for this purpose.
3. There is considerable **practical experience** indicating that second-order models work well in solving real response surface problems.

In general, the first-order model is

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k \quad (1.7)$$

and the second-order model is

$$\eta = \beta_0 + \sum_{j=1}^k \beta_j x_j + \sum_{j=1}^k \beta_{jj} x_j^2 + \sum_{i < j=2}^k \sum_{i=1}^k \beta_{ij} x_i x_j \quad (1.8)$$

In some situations, approximating polynomials of order higher than two are used. The general motivation for a polynomial approximation for the true response function  $f$  is based on the **Taylor series expansion** around the point  $x_{10}, x_{20}, \dots, x_{k0}$ . For example, the first-order model is developed from the first-order Taylor series expansion

$$\begin{aligned} f \cong & f(x_{10}, x_{20}, \dots, x_{k0}) + \left. \frac{\partial f}{\partial x_1} \right|_{\mathbf{x}=\mathbf{x}_0} (x_1 - x_{10}) \\ & + \left. \frac{\partial f}{\partial x_2} \right|_{\mathbf{x}=\mathbf{x}_0} (x_2 - x_{20}) + \cdots + \left. \frac{\partial f}{\partial x_k} \right|_{\mathbf{x}=\mathbf{x}_0} (x_k - x_{k0}) \end{aligned} \quad (1.9)$$

where  $\mathbf{x}$  refers to the vector of independent variables and  $\mathbf{x}_0$  is the vector of independent variables at the specific point  $x_{10}, x_{20}, \dots, x_{k0}$ . In Equation 1.9 we have only included the first-order terms in the expansion, so if we let  $\beta_0 = f(x_{10}, x_{20}, \dots, x_{k0})$ ,  $\beta_1 = (\partial f / \partial x_1)|_{\mathbf{x}=\mathbf{x}_0}$ ,  $\dots$ ,  $\beta_k = (\partial f / \partial x_k)|_{\mathbf{x}=\mathbf{x}_0}$ , we have the first-order approximating model in Equation 1.7. If we were to include second-order terms in Equation 1.9, this would lead to the second-order approximating model in Equation 1.8.

Finally, note that there is a close connection between RSM and **linear regression analysis**. For example, consider the model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

The  $\beta$ 's are a set of unknown parameters. To estimate the values of these parameters, we must collect data on the system we are studying. Regression analysis is a branch of statistical model building that uses these data to estimate the  $\beta$ 's. Because, in general, polynomial models are linear functions of the unknown  $\beta$ 's, we refer to the technique as **linear regression analysis**. We will also see that it is very important to plan the data collection phase of a response surface study carefully. In fact, special types of experimental designs, called **response surface designs**, are valuable in this regard. A substantial part of this book is devoted to response surface designs. Note that analyses and designs need to be carefully matched. If we are planning to analyze data from our planned experiment using a first order model, then the design that we select should be well suited for this analysis. Similarly, if we anticipate curvature similar to what can be modeled with a second-order model, then a different design should be selected.



Good response surface designs have been constructed to perform well based on a particular assumed model, but also have been structured so that they are able to evaluate the assumptions of the model being analyzed to determine if the experimenter's initial impressions of the system under study match the true underlying relationship which produced the data to be analyzed. Hence the experimenter should think carefully about the goals of a particular experiment and what the anticipated analysis will involve before selecting the design for data collection.

### 1.1.2 The Sequential Nature of RSM

Most applications of RSM are **sequential** in nature. That is, at first some ideas are generated concerning which factors or variables are likely to be important in the response surface study. This usually leads to an experiment designed to investigate these factors with a view toward verifying the role of the factors in influencing the response and eliminating the unimportant ones. This type of experiment is usually called a **screening experiment**. Often at the outset of a response surface study there is a rather long list of variables that could be important in explaining the response. The objective of factor screening is to reduce this list of candidate variables to a relative few so that subsequent experiments will be more efficient and require fewer runs or tests. We refer to a screening experiment as **phase zero** of a response surface study. Since interest in a screening experiment lies in understanding the gross behavior of the system and how factors are related to the response, a first-order model is commonly selected. The class of response surface designs which are used for screening experiments are well suited for gaining understanding about the main effects from different independent variables and comparing their relative contributions to changes in the response values. Since this represents an early stage in the planned sequence of experiments, the goal is to determine which of the factors are more influential on the response while using as small a fraction of the total experimental budget as possible. You should never undertake a response surface analysis until a screening experiment has been performed to identify the important factors.

Once the important independent variables are identified, **phase one** of the response surface study begins. In this phase, the experimenter's objective is to determine where the collected data lie relative to an ideal response. Often, there are two possible outcomes with the current levels or settings of the independent variables resulting in a value of the response that is near the optimum (such as the point *B* in Fig. 1.1b), or the process is operating in some other region that is (possibly) remote from the optimum (such as the point *A* in Fig. 1.1b). If the current settings or levels of the independent variables are not consistent with optimum performance, then the experimenter must determine a set of adjustments to the process variables that will move the process toward the optimum. This phase of response surface methodology makes considerable use of the first-order model and an optimization technique called the **method of steepest ascent**. These techniques will be discussed and illustrated in Chapter 5.

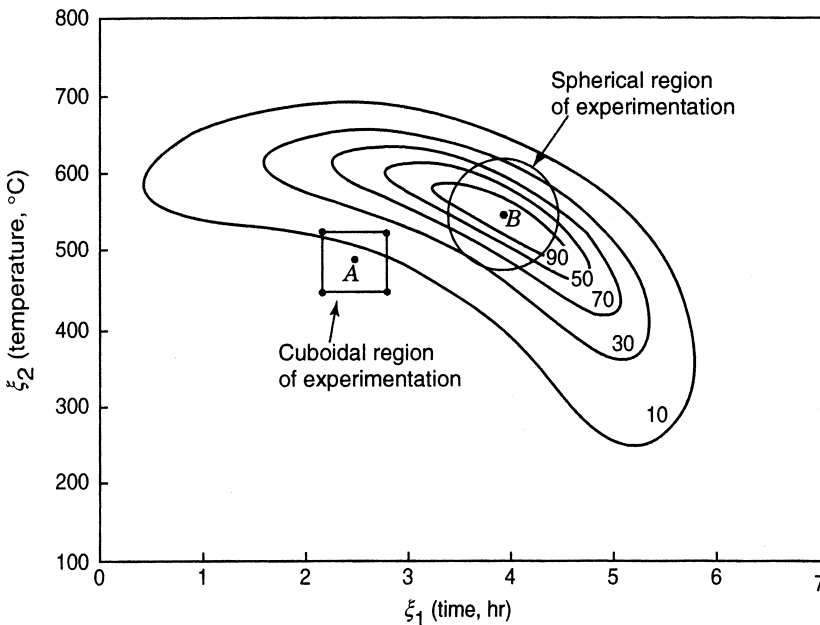
**Phase two** of a response surface study begins when the process is near the optimum. At this point the experimenter usually wants a model that will accurately approximate the true response function within a relatively small region around the optimum. Because the true response surface usually exhibits curvature near the optimum (refer to Fig. 1.1), a second-order model (or very occasionally some higher-order polynomial) will be used. Once an appropriate approximating model has been obtained, this model may be analyzed

to determine the optimum conditions for the process. Chapter 6 will present techniques for the analysis of the second-order model and the determination of optimum conditions.

Response surface designs for modeling the response near the optimum are again selected to match the anticipated analysis. Often, the plan is to characterize the relationship between the response and the key independent variables using the second-order model of the form in Equation 1.8. Designs are constructed to be able to estimate the response for input factor combinations around the expected optimum, where curvature in the relationship is common. Since this stage of experimentation is focused on determining a best set of input values for which the process to operate, a generous portion of the experimental budget is generally reserved for this portion of the process.

A final stage of experimentation, which generally does not require sophisticated response surface designs or a large portion of the experimental budget, is a **confirmatory experiment**. This data collection is generally simple and small, but is designed to confirm that the identified optimum that was obtained in phase two can be achieved by setting the independent variables at the designated settings.

This sequential experimental process is usually performed within some region of the independent variable space called the **operability region**. For the chemical process illustrated in Fig. 1.1, the operability region is  $0 \text{ hr} < \xi_1 \leq 7 \text{ hr}$  and  $100^\circ\text{C} \leq \xi_2 \leq 800^\circ\text{C}$ . Suppose we are currently operating at the levels  $\xi_1 = 2.5 \text{ hr}$  and  $\xi_2 = 500^\circ\text{C}$ , shown as point *A* in Fig. 1.6. Now it is unlikely that we would want to explore the entire region of operability with a single experiment. Instead, we usually define a smaller **region of interest** or **region of experimentation** around the point *A* within the larger region of operability. Typically, this region of experimentation is either a cuboidal region, as shown around the point *A* in Fig. 1.6, or a spherical region, as shown around point *B*. The choice of response surface design matches the specified region of experimentation.



**Figure 1.6** The region of operability and the region of experimentation.

The sequential nature of response surface methodology allows the experimenter to learn about the process or system under study as the investigation proceeds. This ensures that over the course of the RSM application the experimenter will learn the answers to questions such as (1) the location of the region of the optimum, (2) the type of approximating function required, (3) the proper choice of experimental designs, (4) how much replication is necessary, and (5) whether or not transformations on the responses or any of the process variables are required. Because the nature of a response surface study has multiple stages with different goals, there are several important aspects that need to be managed throughout the process. First, many studies have budget constraints that will dictate how much and what data can be collected. It is important to plan for all of the stages of the study and to allow for adequate resources to be available to effectively answer the important questions in each phase.

Second, since the knowledge gained in early phases of the study help to determine what subsequent experiments will study, it is important to plan how the different phases will connect to each other, and what information can be leveraged from early phases. Thirdly, the selection of a model for the analysis of the data from each phase is based on current understanding of the underlying process. It is important to think of the sequence of experiments as a mechanism for not having to make too many assumptions at any stage. Running a large complicated experiment that has many untested assumptions can lead to costly errors and wasting of resources. Hence, a series of smaller experiments can verify some assumptions early in the sequence and can allow the experimenter to proceed in later stages with greater confidence.

Lastly, we again mention the connection between the choice of experiment and the planned analysis. Before jumping in to collect data, the goals of each phase should be clearly defined, and the nature of the response surface design selected should reflect the goals and the planned analysis. Since there are often surprises when collecting and analyzing data, it is helpful to consider what could go wrong with the experiment and to have a plan for how to deal with some of these surprises. A substantial portion of this book—Chapters 3, 4, 8, and 9—is devoted to designed experiments useful in RSM.

### 1.1.3 Objectives and Typical Applications of RSM

Response surface methodology is useful in the solution of many types of industrial problems. Generally, these problems fall into three categories:

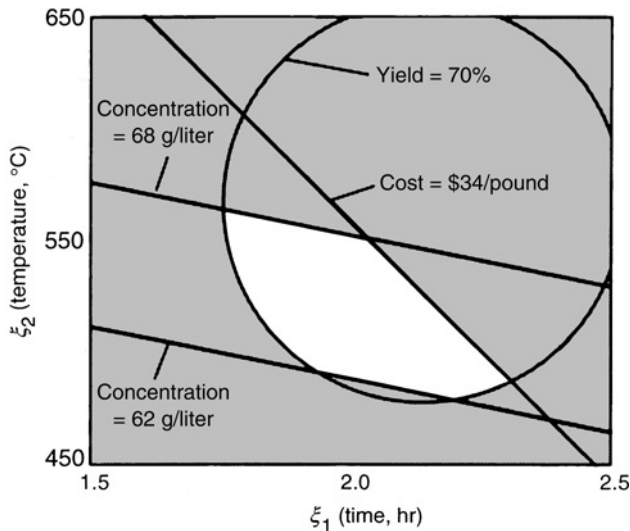
1. *Mapping a Response Surface over a Particular Region of Interest.* Consider the chemical process in Fig. 1.1b. Normally, this process would operate at a particular setting of reaction time and reaction temperature. However, some changes to these normal operating levels might occasionally be necessary, perhaps to produce a product that meets other specific customer requirements. If the true unknown response function has been approximated over a region around the current operating conditions with a suitable fitted response surface (say a second-order surface), then the process engineer can predict in advance the changes in yield that will result from any readjustments to the input variables, namely, time and temperature.
2. *Optimization of the Response.* In the industrial world, a very important problem is determining the conditions that optimize the process. In the chemical process of Fig. 1.1b, this implies determining the levels of time and temperature that result in maximum yield. An RSM study that began near point A in Fig. 1.1b would eventually

lead the experimenter to the region near point  $B$ . A second-order model could then be used to approximate the yield response in a narrow region around point  $B$ , and from examination of this approximating response surface the optimum levels or condition for time and temperature could be chosen.

3. *Selection of Operating Conditions to Achieve Specifications or Customer Requirements.* In most response surface problems there are several responses that must be simultaneously considered. For example, in the chemical process of Fig. 1.1, suppose that in addition to yield, there are two other responses: cost and concentration. We would like to maintain yield above 70%, while simultaneously keeping the cost below \$34/pound; however, the customer has imposed specifications for concentration such that this important physical property must be  $65 \pm 3$  g/liter.

One way that we could solve this problem is to obtain response surfaces for all three responses—yield, cost, and concentration—and then superimpose the contours for these responses in the time–temperature plane, as illustrated in Fig. 1.7. In this figure we have shown the contours for yield = 70%, cost = \$34/pound, concentration = 62 g/liter, and concentration = 68 g/liter. The unshaded region in this figure represents the region containing operating conditions that simultaneously satisfy all requirements on the process.

In practice, complex process optimization problems such as this can often be solved by superimposing appropriate response surface contours. However, it is not unusual to encounter problems with more than two process variables and more complex response requirements to satisfy. In such problems, other optimization methods that are more effective than overlaying contour plots will be necessary, and can often not only identify a region which satisfies the minimal customer requirements, but also find an optimal combination of input variables to achieve ideal performance. We will discuss methodology for solving these types of problems in Chapter 7.



**Figure 1.7** The unshaded region showing the conditions for which yield  $\geq 70\%$ , cost  $\leq \$34/\text{pound}$ , and  $62 \text{ g/liter} \leq \text{concentration} \leq 68 \text{ g/liter}$ .

### 1.1.4 RSM and the Philosophy of Quality Improvement

During the last few decades, industrial organizations in the United States and Europe have become keenly interested in quality and process improvement. Statistical methods, including statistical process control (SPC) and design of experiments, play a key role in this activity. Quality improvement is most effective when it occurs early in the product and process development cycle. It is very difficult, expensive, and inefficient to manufacture a poorly designed product. Industries such as semiconductors and electronics, aerospace, automotive, biotechnology and pharmaceuticals, medical devices, chemical, and process industries are all examples where experimental design methodology has resulted in shorter design and development time for new products, as well as products that are easier to manufacture, have higher reliability, have enhanced field performance, and meet or exceed customer requirements.

RSM is an important branch of experimental design in this regard. RSM is a critical technology in developing new processes, optimizing their performance, and improving the design and/or formulation of new products. It is often an important **concurrent engineering tool**, in that product design, process development, quality, manufacturing engineering, and operations personnel often work together in a team environment to apply RSM. The objectives of quality improvement, including reduction of variability and improved product and process performance, can often be accomplished directly using RSM.

## 1.2 PRODUCT DESIGN AND FORMULATION (MIXTURE PROBLEMS)

Many product design and development activities involve formulation problems, in which two or more ingredients are mixed together. For example, suppose we are developing a new household cleaning product. This product is formulated by mixing several chemical surfactants together. The product engineer or scientist would like to find an appropriate blend of the ingredients so that the grease-cutting capability of the cleaner is good, and so that it generates an appropriate level of foam when in use. In this situation the response variables—namely, grease-cutting ability and amount of foam—depend on the percentages or proportions of the individual chemical surfactants (the ingredients) that are present in the product formulation.

There are many industrial problems where the response variables of interest in the product are a function of the proportions of the different ingredients used in its formulation. This is a special type of response surface problem called a **mixture problem**.

While we traditionally think of mixture problems in the product design or formulation environment, they occur in many other settings. Consider plasma etching of silicon wafers, a common manufacturing process in the semiconductor industry. Etching is usually accomplished by introducing a blend of gases inside a chamber containing the wafers. The measured responses include the etch rate, the uniformity of the etch, and the selectivity (a measure of the relative etch rates of the different materials on the wafer). All of these responses are a function of the proportions of the different ingredients blended together in the etching chamber.

There are special response surface design techniques and model-building methods for mixture problems. These techniques are discussed in Chapters 12 and 13.

### 1.3 ROBUST DESIGN AND PROCESS ROBUSTNESS STUDIES

It is well known that variation in key performance characteristics can result in poor product and process quality. During the 1980s, considerable attention was given to this problem, and methodology was developed for using experimental design, specifically for the following:

1. For designing products or processes so that they are robust to environment conditions.
2. For designing or developing products so that they are robust to component variation.
3. For minimizing variability in the output response of a product around a target value.

By **robust**, we mean that the product or process performs consistently on target and is relatively insensitive to factors that are difficult to control.

Professor Genichi Taguchi used the term **robust parameter design** (or RPD) to describe his approach to this important class of industrial problems. Essentially, robust parameter design methodology strives to reduce product or process variation by choosing levels of controllable factors (or parameters) that make the system insensitive (or robust) to changes in a set of uncontrollable factors that represent most of the sources of variability. Taguchi referred to these uncontrollable factors as **noise factors**. These are the environmental factors such as humidity levels, changes in raw material properties, how the customer will use the product, product aging, and component variability referred to in 1 and 2 above. We usually assume that these noise factors are uncontrollable in the field, but can be controlled during product or process development for purposes of a designed experiment.

Considerable attention has been focused on the methodology advocated by Taguchi, and a number of flaws in his approach have been discovered. However, there are many useful concepts in his philosophy, and it is relatively easy to incorporate these within the framework of response surface methodology. In Chapter 11 we will present the response surface approach to robust design and process robustness studies.

### 1.4 USEFUL REFERENCES ON RSM

The origin of RSM is the seminal paper by Box and Wilson (1951). They also describe the application of RSM to chemical processes. This paper had a profound impact on industrial applications of experimental design, and was the motivation of much of the research in the field. Many of the key research and applications papers are cited in this book.

There have also been five review papers published on RSM: Hill and Hunter (1966), Mead and Pike (1975), Myers et al. (1989), Myers et al. (2004) and Anderson-Cook et al. (2009a). The paper by Myers (1999) on future directions in RSM offers a view of research needs in the field. There are also two other full-length books on the subject: Box and Draper (1987) and Khuri and Cornell (1996). A second edition of the Box and Draper book was published in 2007 with a slightly different title [Box and Draper (2007)]. An edited volume by Khuri (2006) considers some specialized RSM topics. The monograph by Myers (1976) was the first book devoted exclusively to RSM.

---

# 2

---

## BUILDING EMPIRICAL MODELS

### 2.1 LINEAR REGRESSION MODELS

The practical application of response surface methodology (RSM) requires developing an approximating model for the true response surface. The underlying true response surface is typically driven by some unknown **physical mechanism**. The approximating model is based on observed data from the process or system and is an **empirical model**. Multiple regression is a collection of statistical techniques useful for building the types of empirical models required in RSM.

As an example, suppose that we wish to develop an empirical model relating the effective life of a cutting tool to the cutting speed and the tool angle. A first-order response surface model that might describe this relationship is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \quad (2.1)$$

where  $y$  represents the tool life,  $x_1$  represents the cutting speed, and  $x_2$  represents the tool angle. This is a **multiple linear regression model** with two independent variables. We often call the independent variables **predictor variables** or **regressors**. The term “linear” is used because Equation 2.1 is a linear function of the unknown parameters  $\beta_0$ ,  $\beta_1$ , and  $\beta_2$ . The model describes a plane in the two-dimensional  $x_1, x_2$  space. The parameter  $\beta_0$  fixes the intercept of the plane. We sometimes call  $\beta_1$  and  $\beta_2$  **partial regression coefficients**, because  $\beta_1$  measures the expected change in  $y$  per unit change in  $x_1$  when  $x_2$  is held constant, and  $\beta_2$  measures the expected change in  $y$  per unit change in  $x_2$  when  $x_1$  is held constant.

In general, the response variable  $y$  may be related to  $k$  regressor variables. The model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \quad (2.2)$$

is called a **multiple linear regression model** with  $k$  regressor variables. The parameters  $\beta_j, j = 0, 1, \dots, k$ , are called the **regression coefficients**. This model describes a hyperplane in the  $k$ -dimensional space of the regressor variables  $\{x_j\}$ . The parameter  $\beta_j$  represents the expected change in response  $y$  per unit change in  $x_j$  when all the remaining independent variables  $x_i$  ( $i \neq j$ ) are held constant.

Models that are more complex in appearance than Equation 2.2 may often still be analyzed by multiple linear regression techniques. For example, consider adding an **interaction term** to the first-order model in two variables, say

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon \quad (2.3)$$

If we let  $x_3 = x_1 x_2$  and  $\beta_3 = \beta_{12}$ , then Equation 2.3 can be written as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon \quad (2.4)$$

which is a standard multiple linear regression model with three regressors. As another example, consider the **second-order** response surface model in two variables:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \varepsilon \quad (2.5)$$

If we let  $x_3 = x_1^2, x_4 = x_2^2, x_5 = x_1 x_2, \beta_3 = \beta_{11}, \beta_4 = \beta_{22}$ , and  $\beta_5 = \beta_{12}$ , then this becomes

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon \quad (2.6)$$

which is a linear regression model. In general, any regression model that is linear in the parameters (the  $\beta$ -values) is a linear regression model, regardless of the shape of the response surface that it generates.

In this chapter we will present and illustrate methods for estimating the parameters in multiple linear regression models. This is often called **model fitting**. We will also discuss methods for testing hypotheses and constructing confidence intervals for these models, as well as for checking the adequacy of the model fit. Our focus is primarily on those aspects of regression analysis useful in RSM. For more complete presentations of regression, refer to Montgomery, Peck, and Vining (2012) and Myers (1990).

## 2.2 ESTIMATION OF THE PARAMETERS IN LINEAR REGRESSION MODELS

The method of least squares is typically used to estimate the regression coefficients in a multiple linear regression model. Suppose that  $n > k$  observations on the response variable are available, say  $y_1, y_2, \dots, y_n$ . Along with each observed response  $y_i$ , we will have an observation on each regressor variable, and let  $x_{ij}$  denote the  $i$ th observation or level of variable  $x_j$ . Table 2.1 shows the data layout. We assume that the error term  $\varepsilon$  in the model has  $E(\varepsilon) = 0$  and  $\text{Var}(\varepsilon) = \sigma^2$  and that the  $\{\varepsilon_i\}$  are uncorrelated random variables.



**TABLE 2.1 Data for Multiple Linear Regression**

$y$	$x_1$	$x_2$	$\dots$	$x_k$
$y_1$	$x_{11}$	$x_{12}$	$\dots$	$x_{1k}$
$y_2$	$x_{21}$	$x_{22}$	$\dots$	$x_{2k}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$y_n$	$x_{n1}$	$x_{n2}$	$\dots$	$x_{nk}$

We may write the model equation (Eq. 2.2) in terms of the observations in Table 2.1 as

$$\begin{aligned}
 y_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i \\
 &= \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \varepsilon_i, \quad i = 1, 2, \dots, n
 \end{aligned}
 \tag{2.7}$$

The method of least squares chooses the  $\beta$ 's in Equation 2.7 so that the sum of the squares of the errors,  $\varepsilon_i$ , are minimized. The least squares function is

$$\begin{aligned}
 L &= \sum_{i=1}^n \varepsilon_i^2 \\
 &= \sum_{i=1}^n \left( y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{ij} \right)^2
 \end{aligned}
 \tag{2.8}$$

The function  $L$  is to be minimized with respect to  $\beta_0, \beta_1, \dots, \beta_k$ . The **least squares estimators**, say  $b_0, b_1, \dots, b_k$ , must satisfy

$$\left. \frac{\partial L}{\partial \beta_0} \right|_{b_0, b_1, \dots, b_k} = -2 \sum_{i=1}^n \left( y_i - b_0 - \sum_{j=1}^k b_j x_{ij} \right) = 0
 \tag{2.9a}$$

and

$$\left. \frac{\partial L}{\partial \beta_j} \right|_{b_0, b_1, \dots, b_k} = -2 \sum_{i=1}^n \left( y_i - b_0 - \sum_{j=1}^k b_j x_{ij} \right) x_{ij} = 0,
 \tag{2.9b}$$

where  $j = 1, 2, \dots, k$ . Simplifying Equation 2.9, we obtain

$$\begin{aligned}
 nb_0 + b_1 \sum_{i=1}^n x_{i1} + b_2 \sum_{i=1}^n x_{i2} + \dots + b_k \sum_{i=1}^n x_{ik} &= \sum_{i=1}^n y_i \\
 b_0 \sum_{i=1}^n x_{i1} + b_1 \sum_{i=1}^n x_{i1}^2 + b_2 \sum_{i=1}^n x_{i1}x_{i2} + \dots + b_k \sum_{i=1}^n x_{i1}x_{ik} &= \sum_{i=1}^n x_{i1}y_i \\
 b_0 \sum_{i=1}^n x_{ik} + b_1 \sum_{i=1}^n x_{ik}x_{i1} + b_2 \sum_{i=1}^n x_{ik}x_{i2} + \dots + b_k \sum_{i=1}^n x_{ik}^2 &= \sum_{i=1}^n x_{ik}y_i
 \end{aligned}
 \tag{2.10}$$

These equations are called the **least squares normal equations**. Note that there are  $p = k + 1$  normal equations, one for each of the unknown regression coefficients. The solution to the normal equations will be the least squares estimators of the regression coefficients  $b_0, b_1, \dots, b_k$ .

It is simpler to solve the normal equations if they are expressed in matrix notation. We now give a matrix development of the normal equations that parallels the development of Equation 2.10. The model in terms of the observations, Equation 2.7, may be written in matrix notation as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \text{and} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

In general,  $\mathbf{y}$  is an  $n \times 1$  vector of the observations,  $\mathbf{X}$  is an  $n \times p$  **model matrix** consisting of the levels of the independent variables expanded to model form,  $\boldsymbol{\beta}$  is a  $p \times 1$  vector of the regression coefficients, and  $\boldsymbol{\varepsilon}$  is an  $n \times 1$  vector of random errors. Notice that  $\mathbf{X}$  consists of the columns containing the independent variables from Table 2.1 plus an additional column of 1s to account for the intercept term in the model.

We wish to find the vector of least squares estimators,  $\mathbf{b}$ , that minimizes

$$L = \sum_{i=1}^n \varepsilon_i^2 = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

Note that  $L$  may be expressed as

$$\begin{aligned} L &= \mathbf{y}'\mathbf{y} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \\ &= \mathbf{y}'\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \end{aligned} \quad (2.11)$$

since  $\boldsymbol{\beta}'\mathbf{X}'\mathbf{y}$  is a  $1 \times 1$  matrix, or a scalar, and its transpose  $(\boldsymbol{\beta}'\mathbf{X}'\mathbf{y})' = \mathbf{y}'\mathbf{X}\boldsymbol{\beta}$  is the same scalar. The least squares estimators must satisfy

$$\left. \frac{\partial L}{\partial \boldsymbol{\beta}} \right|_{\mathbf{b}} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{0}$$

which simplifies to

$$\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{y} \quad (2.12)$$

Equation 2.12 is the set of least squares normal equations in matrix form. It is identical to Equation 2.10. To solve the normal equations, multiply both sides of Equation 2.12 by the inverse of  $\mathbf{X}'\mathbf{X}$ . Thus, the least squares estimator of  $\boldsymbol{\beta}$  is

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (2.13)$$

It is easy to see that the matrix form of the normal equations is identical to the scalar form. Writing out Equation 2.12 in detail, we obtain

$$\begin{bmatrix} n & \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i2} & \cdots & \sum_{i=1}^n x_{ik} \\ \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i1}^2 & \sum_{i=1}^n x_{i1}x_{i2} & \cdots & \sum_{i=1}^n x_{i1}x_{ik} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n x_{ik} & \sum_{i=1}^n x_{ik}x_{i1} & \sum_{i=1}^n x_{ik}x_{i2} & \cdots & \sum_{i=1}^n x_{ik}^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_{i1}y_i \\ \vdots \\ \sum_{i=1}^n x_{ik}y_i \end{bmatrix}$$

If the indicated matrix multiplication is performed, the scalar form of the normal equations (i.e., Eq. 2.10) will result. In this form it is easy to see that  $\mathbf{X}'\mathbf{X}$  is a  $p \times p$  symmetric matrix and  $\mathbf{X}'\mathbf{y}$  is a  $p \times 1$  column vector. Note the special structure of the matrix  $\mathbf{X}'\mathbf{X}$ . The diagonal elements of  $\mathbf{X}'\mathbf{X}$  are the sums of squares of the elements in the columns of  $\mathbf{X}$ , and the off-diagonal elements are the sums of cross-products of the elements in the columns of  $\mathbf{X}$ . Furthermore, note that the elements of  $\mathbf{X}'\mathbf{y}$  are the sums of cross-products of the columns of  $\mathbf{X}$  and the observations  $\{y_i\}$ .

The fitted regression model is

$$\hat{y} = \mathbf{X}\mathbf{b} \quad (2.14)$$

In scalar notation, the fitted model is

$$\hat{y}_i = b_0 + \sum_{j=1}^k b_j x_{ij}, \quad i = 1, 2, \dots, n$$

The difference between the observation  $y_i$  and the fitted value  $\hat{y}_i$  is a **residual**, say  $e_i = y_i - \hat{y}_i$ . The  $n \times 1$  vector of residuals is denoted by

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} \quad (2.15)$$

**Example 2.1 The Transistor Gain Data** The transistor gain in an integrated circuit device between emitter and collector ( $hFE$ ) is related to two variables that can be controlled at the deposition process, emitter drive-in time ( $\xi_1$ , in minutes), and emitter dose ( $\xi_2$ , units of  $10^{14}$  ions). Fourteen samples were observed following deposition, and the resulting data are shown in Table 2.2. We will fit a linear regression model using gain as the response and emitter drive-in time and emitter dose as the regressor variables.

**TABLE 2.2 Data on Transistor Gain ( $y$ ) for Example 2.1**

Observation	$\xi_1$ (drive-in time, minutes)	$\xi_2$ (dose, $10^{14}$ ions)	$x_1$	$x_2$	$y$ (gain or $hFE$ )
1	195	4.00	-1	-1	1004
2	255	4.00	1	-1	1636
3	195	4.60	-1	0.6667	852
4	255	4.60	1	0.6667	1506
5	225	4.20	0	-0.4444	1272
6	225	4.10	0	-0.7222	1270
7	225	4.60	0	0.6667	1269
8	195	4.30	-1	-0.1667	903
9	255	4.30	1	-0.1667	1555
10	225	4.00	0	-1	1260
11	225	4.70	0	0.9444	1146
12	225	4.30	0	-0.1667	1276
13	225	4.72	0	1	1225
14	230	4.30	0.1667	-0.1667	1321

Columns 2 and 3 of Table 2.2 show the actual or natural unit values of  $\xi_1$  and  $\xi_2$ , while columns 4 and 5 contain values of the corresponding coded variables  $x_1$  and  $x_2$ , where

$$x_{i1} = \frac{\xi_{i1} - [\max(\xi_{i1}) + \min(\xi_{i1})]/2}{[\max(\xi_{i1}) - \min(\xi_{i1})]/2} = \frac{\xi_{i1} - (255 + 195)/2}{(255 - 195)/2} = \frac{\xi_{i1} - 225}{30}$$

$$x_{i2} = \frac{\xi_{i2} - [\max(\xi_{i2}) + \min(\xi_{i2})]/2}{[\max(\xi_{i2}) - \min(\xi_{i2})]/2} = \frac{\xi_{i2} - (4.72 + 4.00)/2}{(4.72 - 4.00)/2} = \frac{\xi_{i2} - 4.36}{0.36}$$

This coding scheme is widely used in fitting linear regression models, and it results in all the values of  $x_1$  and  $x_2$  falling between  $-1$  and  $+1$ , as shown in Table 2.2.

We will fit the model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

using the coded variables. The model matrix  $\mathbf{X}$  and vector  $\mathbf{y}$  are

$$\mathbf{X} = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 0.6667 \\ 1 & 1 & 0.6667 \\ 1 & 0 & -0.4444 \\ 1 & 0 & -0.7222 \\ 1 & 0 & 0.6667 \\ 1 & -1 & -0.1667 \\ 1 & 1 & -0.1667 \\ 1 & 0 & -1 \\ 1 & 0 & 0.9444 \\ 1 & 0 & -0.1667 \\ 1 & 0 & 1 \\ 1 & 0.1667 & -0.1667 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1004 \\ 1636 \\ 852 \\ 1506 \\ 1272 \\ 1270 \\ 1269 \\ 903 \\ 1555 \\ 1260 \\ 1146 \\ 1276 \\ 1225 \\ 1321 \end{bmatrix}$$

The matrix  $\mathbf{X}'\mathbf{X}$  is

$$\begin{aligned}\mathbf{X}'\mathbf{X} &= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ -1 & 1 & \cdots & 0.1667 \\ -1 & -1 & \cdots & -0.1667 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ \vdots & \vdots & \vdots \\ 1 & 0.1667 & -0.1667 \end{bmatrix} \\ &= \begin{bmatrix} 14 & 0.1667 & -0.8889 \\ 0.1667 & 6.027789 & -0.02779 \\ -0.8889 & -0.02779 & 7.055578 \end{bmatrix}\end{aligned}$$

and the vector  $\mathbf{X}'\mathbf{y}$  is

$$\begin{aligned}\mathbf{X}'\mathbf{y} &= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ -1 & 1 & \cdots & 0.1667 \\ -1 & -1 & \cdots & -0.1667 \end{bmatrix} \begin{bmatrix} 1004 \\ 1636 \\ \vdots \\ 1321 \end{bmatrix} \\ &= \begin{bmatrix} 17,495 \\ 2158.211 \\ -1499.74 \end{bmatrix}\end{aligned}$$

The least squares estimate of  $\beta$  is

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

or

$$\begin{aligned}\mathbf{b} &= \begin{bmatrix} 0.072027 & -0.00195 & 0.009067 \\ -0.00195 & 0.165954 & 0.000408 \\ 0.009067 & 0.000408 & 0.142876 \end{bmatrix} \begin{bmatrix} 17,495 \\ 2158.211 \\ -1499.74 \end{bmatrix} \\ &= \begin{bmatrix} 1242.3057 \\ 323.4366 \\ -54.7691 \end{bmatrix}\end{aligned}$$

The least squares fit with the regression coefficients reported to one decimal place is

$$\hat{y} = 1242.3 + 323.4x_1 - 54.8x_2$$

This can be converted into an equation using the natural variables  $\xi_1$  and  $\xi_2$  by substituting the relationships between  $x_1$  and  $\xi_1$  and  $x_2$  and  $\xi_2$  as follows:

$$\hat{y} = 1242.3 + 323.4 \left( \frac{\xi_1 - 225}{30} \right) - 54.8 \left( \frac{\xi_2 - 4.36}{0.36} \right)$$

or

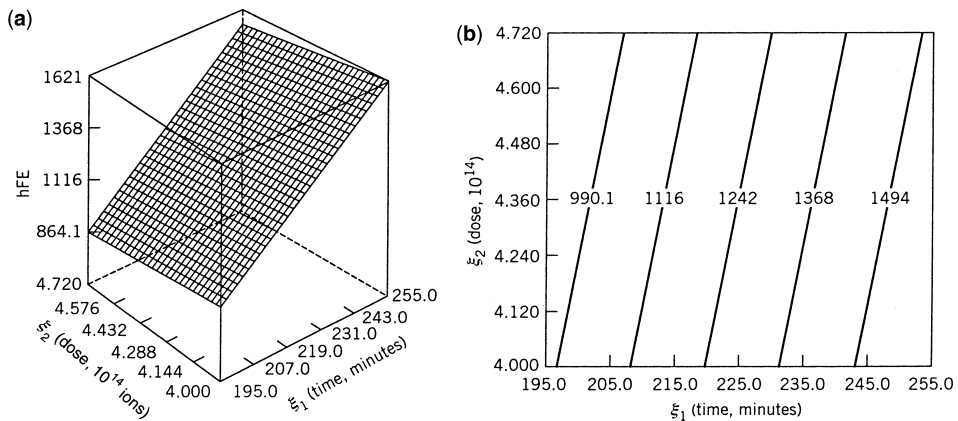
$$\hat{y} = -520.1 + 10.781\xi_1 - 152.15\xi_2$$

**TABLE 2.3 Observations, Fitted Values, Residuals, and Other Summary Information for Example 2.1**

Observation	$y_i$	$\hat{y}_{ii}$	$e_i$	$h_{ii}$	$r_i$	$t_i$	$D_i$
1	1004.0	973.7	30.3	0.367	1.092	1.103	0.231
2	1636.0	1620.5	15.5	0.358	0.553	0.535	0.057
3	852.0	882.4	-30.4	0.317	-1.052	-1.057	0.171
4	1506.0	1529.2	-23.2	0.310	-0.801	-0.787	0.096
5	1272.0	1266.7	5.3	0.092	0.160	0.153	0.001
6	1270.0	1281.9	-11.9	0.133	-0.365	-0.350	0.007
7	1269.0	1205.8	63.2	0.148	1.960	2.316	0.222
8	903.0	928.0	-25.0	0.243	-0.823	-0.810	0.072
9	1555.0	1574.9	-19.9	0.235	-0.651	-0.633	0.043
10	1260.0	1297.1	-37.1	0.197	-1.185	-1.209	0.115
11	1146.0	1190.6	-44.6	0.217	-1.442	-1.527	0.192
12	1276.0	1251.4	24.6	0.073	0.730	0.714	0.014
13	1225.0	1187.5	37.5	0.233	1.225	1.256	0.152
14	1321.0	1305.3	15.7	0.077	0.466	0.449	0.006

Table 2.3 shows the observed values of  $y_i$ , the corresponding fitted values  $\hat{y}_i$ , and the residuals from this model. There are several other quantities given in this table that will be defined and discussed later. Figure 2.1 shows the fitted response surface and the contour plot for this model. The response surface for gain is a plane lying above the time–dose space.

**Computing** Statistics software is usually employed to fit regression models. Table 2.4 and Fig. 2.2 present some of the output for the transistor gain data in Example 2.1 from JMP, a widely-used software package that supports regression, experimental design, and RSM. This model was fit to the coded variables in Table 2.2. The first portion of the display

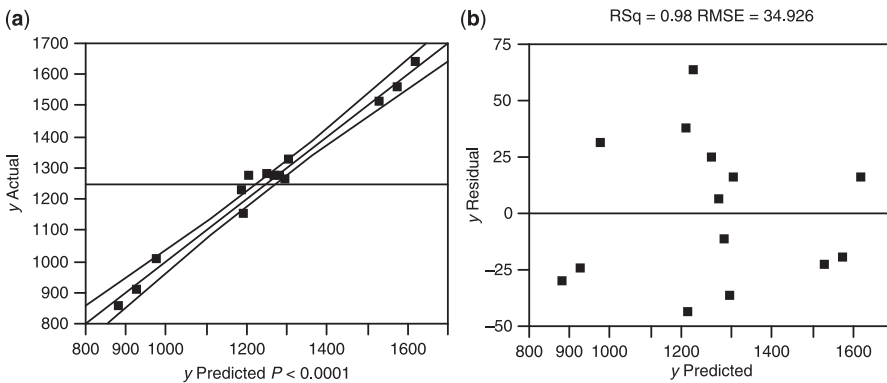


**Figure 2.1** (a) Response surface for gain, Example 2.1. (b) The gain contour plot.

**TABLE 2.4 Regression Output From JMP**

<i>Summary of Fit</i>				
<i>R</i> -Square		0.979835		
<i>R</i> -Square Adj		0.976168		
Root Mean Square Error		34.92553		
Mean of Response		1249.643		
Observations (or Sum Wgts)		14		
<i>Analysis of Variance</i>				
Source	DF	Sum of Squares	Mean Square	<i>F</i> -Ratio
Model	2	651969.49	325985	267.2460
Error	11	13417.72	1220	Prob > <i>F</i>
C. Total	13	665387.21		<0.0001
<i>Parameter Estimates</i>				
Term	Estimate	Std Error	<i>t</i> -Ratio	Prob >   <i>t</i>
Intercept	1242.3181	9.373196	132.54	<0.0001
<i>x</i> 1	323.4253	14.22778	22.73	<0.0001
<i>x</i> 2	-54.77165	13.2001	-4.15	0.0016

in Fig. 2.2 is a plot of the values of the observed response  $y$  versus the predicted values  $\hat{y}_i$  (see Fig. 2.2a). The pairs  $(y_i, \hat{y}_i)$  lie closely along a straight line (the straight line in the graph is a result of a least squares fit). This is usually a good indication that the model is a satisfactory fit to the data. We will discuss other checks of model adequacy later in this chapter. Notice that the estimates of the regression coefficients closely match those that we have computed manually (it is not unusual to find minor differences between manual and computer software regression calculations because of round-off). In subsequent sections we will show how some of the other quantities in the output are obtained and how to interpret them.



**Figure 2.2** Regression output from JMP. (a) Response  $y$  whole model, actual by predicted plot. (b) Residual by predicted plot.

### 2.3 PROPERTIES OF THE LEAST SQUARES ESTIMATORS AND ESTIMATION OF $\sigma^2$

The method of least squares produces an **unbiased estimator** of the parameter  $\beta$  in the multiple linear regression model. This property may be easily demonstrated by finding the expected value of  $\mathbf{b}$  as follows:

$$\begin{aligned} E(\mathbf{b}) &= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}] \\ &= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\beta + \varepsilon)] \\ &= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\varepsilon] \\ &= \beta \end{aligned}$$

because  $E(\varepsilon) = \mathbf{0}$  and  $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X} = \mathbf{I}$ . Thus  $\mathbf{b}$  is an unbiased estimator of  $\beta$ . The variance property of  $\mathbf{b}$  is expressed by the covariance matrix

$$\text{Cov}(\mathbf{b}) = E\{[\mathbf{b} - E(\mathbf{b})][\mathbf{b} - E(\mathbf{b})]'\}$$

The covariance matrix of  $\mathbf{b}$  is a  $p \times p$  symmetric matrix whose  $(j, j)$ th element is the variance of  $b_j$  and whose  $(i, j)$ th element is the covariance between  $b_i$  and  $b_j$ . The **covariance matrix** of  $\mathbf{b}$  is

$$\text{Cov}(\mathbf{b}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \quad (2.16)$$

It is also usually necessary to estimate  $\sigma^2$ . To develop an estimator of this parameter consider the sum of squares of the residuals, say

$$\begin{aligned} SS_E &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 \\ &= \sum_{i=1}^n e_i^2 \\ &= \mathbf{e}'\mathbf{e} \end{aligned}$$

Substituting  $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\mathbf{b}$ , we have

$$\begin{aligned} SS_E &= (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b}) \\ &= \mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}\mathbf{b} + \mathbf{b}'\mathbf{X}'\mathbf{X}\mathbf{b} \\ &= \mathbf{y}'\mathbf{y} - 2\mathbf{b}'\mathbf{X}'\mathbf{y} + \mathbf{b}'\mathbf{X}'\mathbf{X}\mathbf{b} \end{aligned}$$

Because  $\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{y}$ , this last equation becomes

$$SS_E = \mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y} \quad (2.17)$$

Equation 2.17 is called the **error** or **residual sum of squares**, and it has  $n - p$  degrees of freedom associated with it. It can be shown that

$$E(SS_E) = \sigma^2(n - p)$$



so an unbiased estimator of  $\sigma^2$  is given by

$$\hat{\sigma}^2 = \frac{SS_E}{n - p} \quad (2.18)$$

**Example 2.2 The Transistor Gain Data** We will estimate  $\sigma^2$  for the regression model for the transistor gain data from Example 2.1. Because

$$\mathbf{y}'\mathbf{y} = \sum_{i=1}^{14} y_i^2 = 22,527,889.0$$

and

$$\mathbf{b}'\mathbf{X}'\mathbf{y} = \begin{bmatrix} 1242.3 & 323.4 & -54.8 \end{bmatrix} \begin{bmatrix} 17,495 \\ 2158.211 \\ -1499.74 \end{bmatrix} = 22,514,467.9$$

the residual sum of squares is

$$\begin{aligned} SS_E &= \mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y} \\ &= 22,527,889.0 - 22,514,467.9 \\ &= 13,421.1 \end{aligned}$$

Therefore, the estimate of  $\sigma^2$  is computed from Equation 2.18 as follows:

$$\hat{\sigma}^2 = \frac{SS_E}{n - p} = \frac{13,421.1}{14 - 3} = 1220.1$$

Notice that the JMP output in Table 2.4 computes the residual sum of squares (look under the analysis of variance section of the output) as 13,417.72. The difference between the two values is round-off. Both the manual calculations and JMP produce virtually identical estimates of  $\sigma^2$ .

The estimate of  $\sigma^2$  produced by Equation 2.18 is **model-dependent**. That is, it depends on the **form** of the model that is fit to the data. To illustrate this point, suppose that we fit a quadratic model to the gain data, say

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \varepsilon$$

In this model it can be shown that  $SS_E = 12,479.8$ . Because the number of model parameters,  $p$ , equals 6, the estimate of  $\sigma^2$  based on this model is

$$\hat{\sigma}^2 = \frac{12,479.8}{14 - 6} = 1559.975$$

This estimate of  $\sigma^2$  is actually larger than the estimate obtained from the first-order model, suggesting that the first-order model is superior to the quadratic in that there is less unexplained variability resulting from the first-order fit. If replicate runs are available (that is, more than one observation on  $y$  at the same  $x$ -levels), then a **model-independent** estimate of  $\sigma^2$  can be obtained. We will show how to do this in Section 2.7.4.

## 2.4 HYPOTHESIS TESTING IN MULTIPLE REGRESSION

In multiple linear regression problems, certain tests of hypotheses about the model parameters are helpful in measuring the usefulness of the model. In this section, we describe several important hypothesis-testing procedures. These procedures require that the errors  $\varepsilon_i$  in the model be normally and independently distributed with mean zero and variance  $\sigma^2$ , abbreviated  $\varepsilon \sim \text{NID}(0, \sigma^2)$ . As a result of this assumption, the observations  $y_i$  are normally and independently distributed with mean  $\beta_0 + \sum_{j=1}^k \beta_j x_{ij}$  and variance  $\sigma^2$ .

### 2.4.1 Test for Significance of Regression

The test for significance of regression is a test to determine if there is a linear relationship between the response variable  $y$  and a subset of the regressor variables  $x_1, x_2, \dots, x_k$ . The appropriate hypotheses are

$$\begin{aligned} H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0 \\ H_1: \beta_j \neq 0 \text{ for at least one } j \end{aligned} \quad (2.19)$$

Rejection of  $H_0$  in Equation 2.19 implies that at least one of the regressor variables  $x_1, x_2, \dots, x_k$  contributes significantly to the model. The test procedure involves partitioning the total sum of squares  $SS_T = \sum_{i=1}^n (y_i - \bar{y})^2$  into a sum of squares due to the model (or to regression) and a sum of squares due to residual (or error), say

$$SS_T = SS_R + SS_E \quad (2.20)$$

Now if the null hypothesis  $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$  is true, then  $SS_R/\sigma^2$  is distributed as  $\chi_k^2$ , where the number of degrees of freedom for  $\chi^2$  is equal to the number of regressor variables in the model. Also, we can show that  $SS_E/\sigma^2$  is distributed as  $\chi_{n-k-1}^2$  and that  $SS_E$  and  $SS_R$  are independent. The test procedure for  $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$  is to compute

$$F_0 = \frac{SS_R/k}{SS_E/(n-k-1)} = \frac{MS_R}{MS_E} \quad (2.21)$$

and to reject  $H_0$  if  $F_0$  exceeds  $F_{\alpha, k, n-k-1}$ . Alternatively, one could use the  $P$ -value approach to hypothesis testing and, thus, reject  $H_0$  if the  $P$ -value for the statistic  $F_0$  is less than  $\alpha$ . The test is usually summarized in a table such as Table 2.5. This test procedure is called an **analysis of variance (ANOVA)** because it is based on a decomposition of the total variability in the response variable  $y$ .

A computational formula for  $SS_R$  may be found easily. We have derived a computational formula for  $SS_E$  in Equation 2.17—that is,

$$SS_E = \mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y}$$

**TABLE 2.5 Analysis of Variance for Significance of Regression in Multiple Regression**

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	$F_0$
Regression	$SS_R$	$k$	$MS_R$	$MS_R/MS_E$
Error or residual	$SS_E$	$n - k - 1$	$MS_E$	
Total	$SS_T$	$n - 1$		

Now because  $SS_T = \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2 / n = \mathbf{y}'\mathbf{y} - (\sum_{i=1}^n y_i)^2 / n$ , we may rewrite the foregoing equation as

$$SS_E = \mathbf{y}'\mathbf{y} - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n} - \left[ \mathbf{b}'\mathbf{X}'\mathbf{y} - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n} \right]$$

or

$$SS_E = SS_T - SS_R$$

Therefore, the regression sum of squares is

$$SS_R = \mathbf{b}'\mathbf{X}'\mathbf{y} - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n} \quad (2.22)$$

the error sum of squares is

$$SS_E = \mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y} \quad (2.23)$$

and the total sum of squares is

$$SS_T = \mathbf{y}'\mathbf{y} - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n} \quad (2.24)$$

**Example 2.3 The Transistor Gain Data** We will test for significance of regression using the model fit to the transistor gain data for Example 2.1. Note that

$$\begin{aligned} SS_T &= \mathbf{y}'\mathbf{y} - \frac{\left(\sum_{i=1}^{14} y_i\right)^2}{14} \\ &= 22,527,889.0 - \frac{(17,495)^2}{14} \\ &= 665,387.2 \end{aligned}$$

$$\begin{aligned} SS_R &= \mathbf{b}'\mathbf{X}'\mathbf{y} - \frac{\left(\sum_{i=1}^{14} y_i\right)^2}{14} \\ &= 22,514,467.9 - 21,862,501.8 \\ &= 651,966.1 \end{aligned}$$

**TABLE 2.6 Test for Significance of Regression, Example 2.3**

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	$F_0$	$P$ -Value
Regression	651,996.1	2	325,983.0	267.2	$4.74 \times 10^{-10}$
Error	13,421.1	11	1220.1		
Total	665,387.2	13			

and

$$\begin{aligned} SS_E &= SS_T - SS_R \\ &= 665,387.2 - 651,966.1 \\ &= 13,421.1 \end{aligned}$$

The analysis of variance is shown in Table 2.6. If we select  $\alpha = 0.05$ , then we reject  $H_0: \beta_1 = \beta_2 = 0$  because  $F_0 = 267.2 > F_{0.05,2,11} = 3.98$ . Also, note that the  $P$ -value for  $F_0$  (shown in Table 2.6) is considerably smaller than  $\alpha = 0.05$ . Finally, compare the ANOVA in Table 2.6 to the one in the JMP output in Table 2.4. There is a small difference in the “regression sum of squares” in the two tables. This is due to round-off again.

The coefficient of multiple determination  $R^2$  is defined as

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_E}{SS_T} \quad (2.25)$$

$R^2$  is a measure of the amount of reduction in the variability of  $y$  obtained by using the regressor variables  $x_1, x_2, \dots, x_k$  in the model. From inspection of the analysis of the variance identity equation (Eq. 2.20) we see that  $0 \leq R^2 \leq 1$ . However, a large value of  $R^2$  does not necessarily imply that the regression model is good one. Adding a variable to the model will always increase  $R^2$ , regardless of whether the additional variable is statistically significant or not. Thus it is possible for models that have large values of  $R^2$  to yield poor predictions of new observations or estimates of the mean response.

To illustrate, consider the first-order model for the transistor gain data. The value of  $R^2$  for this model is

$$R^2 = \frac{SS_R}{SS_T} = \frac{651,966.1}{665,387.2} = 0.9798$$

That is, the first-order model explains about 97.98% of the variability observed in the gain. Now if we add quadratic terms to this model, we can show that the value of  $R^2$  increases to 0.9812. This increase in  $R^2$  is relatively small, suggesting that the quadratic terms do not really improve the model. The JMP output in Table 2.4 reports the  $R^2$  statistic.

Because  $R^2$  always increases as we add terms to the model, some regression model builders prefer to use an **adjusted**  $R^2$  statistic defined as

$$R_{\text{adj}}^2 = 1 - \frac{SS_E/(n-p)}{SS_T/(n-1)} = 1 - \frac{n-1}{n-p}(1-R^2) \quad (2.26)$$

In general, the adjusted  $R^2$  statistic will not always increase as variables are added to the model. In fact, if unnecessary terms are added, the value of  $R^2_{\text{adj}}$  will often decrease.

For example, consider the transistor gain data. The adjusted  $R^2$  for the first-order model is

$$\begin{aligned} R^2_{\text{adj}} &= 1 - \frac{n-1}{n-p}(1 - R^2) \\ &= 1 - \frac{13}{11}(1 - 0.9798) \\ &= 0.9762 \end{aligned}$$

which is very close to the ordinary  $R^2$  for the first-order model (JMP reports the adjusted  $R^2$  statistic—refer to Table 2.4). When  $R^2$  and  $R^2_{\text{adj}}$  differ dramatically, there is a good chance that nonsignificant terms have been included in the model. Now when the quadratic terms are added to the first-order model, we can show that  $R^2_{\text{adj}} = 0.9695$ ; that is, the adjusted  $R^2$  actually decreases when the quadratic terms are included in the model. This is a strong indication that the quadratic terms are unnecessary.

#### 2.4.2 Tests on Individual Regression Coefficients and Groups of Coefficients

**Individual Regression Coefficients** We are frequently interested in testing hypotheses on the individual regression coefficients. Such tests would be useful in determining the value of each of the regressor variables in the regression model. For example, the model might be more effective with the inclusion of additional variables, or perhaps with the deletion of one or more of the variables already in the model.

Adding a variable to the regression model always causes the sum of squares for regression to increase and the error sum of squares to decrease. We must decide whether the increase in the regression sum of squares is sufficient to warrant using the additional variable in the model. Furthermore, adding an unimportant variable to the model can actually increase the mean square error, thereby decreasing the usefulness of the model.

The hypotheses for testing the significance of any individual regression coefficient, say  $\beta_j$ , are

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0$$

If  $H_0: \beta_j = 0$  is not rejected, then this indicates that  $x_j$  can be deleted from the model. The test statistic for this hypothesis is

$$t_0 = \frac{b_j}{\sqrt{\hat{\sigma}^2 C_{jj}}} \quad (2.27)$$

where  $C_{jj}$  is the diagonal element of  $(\mathbf{X}'\mathbf{X})^{-1}$  corresponding to  $b_j$ . The null hypothesis  $H_0: \beta_j = 0$  is rejected if  $|t_0| > t_{\alpha/2, n-k-1}$ . A  $P$ -value approach could also be used. See Montgomery et al. (2012) for more details. Note that this is really a partial or marginal test, because the

regression coefficient  $b_j$  depends on all the other regressor variables  $x_i$  ( $i \neq j$ ) that are in the model.

The denominator of Equation 2.27,  $\sqrt{\hat{\sigma}^2 C_{jj}}$ , is often called the **standard error** of the regression coefficient  $b_j$ . That is,

$$se(b_j) = \sqrt{\hat{\sigma}^2 C_{jj}} \quad (2.28)$$

Therefore, an equivalent way to write the test statistic in Equation 2.27 is

$$t_0 = \frac{b_j}{se(b_j)} \quad (2.29)$$

**Example 2.4 Tests on Individual Regression Coefficients for the Transistor Gain Data** Consider the regression model for the transistor gain data. We will construct the  $t$ -statistic for the hypotheses  $H_0: \beta_1 = 0$  and  $H_0: \beta_2 = 0$ . The main diagonal elements of  $(\mathbf{X}'\mathbf{X})^{-1}$  corresponding to  $\beta_1$  and  $\beta_2$  are  $C_{11} = 0.165954$  and  $C_{22} = 0.142876$ , respectively, so the two  $t$ -statistics are computed as follows:

$$\begin{aligned} \text{For } H_0: \beta_1 = 0: \quad t_0 &= \frac{b_1}{\sqrt{\hat{\sigma}^2 C_{11}}} \\ &= \frac{323.4}{\sqrt{(1220.1)(0.165954)}} = \frac{323.4}{14.2} = 22.73 \\ \text{For } H_0: \beta_2 = 0: \quad t_0 &= \frac{b_2}{\sqrt{\hat{\sigma}^2 C_{22}}} = \frac{-54.8}{\sqrt{(1220.1)(0.142876)}} \\ &= \frac{-54.8}{13.2} = -4.15 \end{aligned}$$

The absolute values of these  $t$ -statistics are compared with  $t_{0.025,11} = 2.201$  (assuming that we select  $\alpha = 0.05$ ). The magnitudes of both of the  $t$ -statistics are larger than this criterion. Consequently we conclude that  $\beta_1 \neq 0$ , which implies that  $x_1$  contributes significantly to the model given that  $x_2$  is included, and that  $\beta_2 \neq 0$ , which implies that  $x_2$  contributes significantly to the model given that  $x_1$  is included. JMP (refer to Table 2.4) provides the  $t$ -statistics for each regression coefficient. A  $P$ -value approach is used for statistical testing. The  $P$ -values for the two test statistics reported above are small (less than 0.05, say) indicating that both variables contribute significantly to the model.

**Tests on Groups of Coefficients** We may also directly examine the contribution to the regression sum of squares for a particular variable, say  $x_j$ , given that other variables  $x_i$  ( $i \neq j$ ) are included in the model. The procedure used to do this is called the **extra sum of squares method**. This procedure can also be used to investigate the contribution of a *subset* of the regressor variables to the model. Consider the regression model with  $k$  regressor variables:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where  $\mathbf{y}$  is  $n \times 1$ ,  $\mathbf{X}$  is  $n \times p$ ,  $\boldsymbol{\beta}$  is  $p \times 1$ ,  $\boldsymbol{\varepsilon}$  is  $n \times 1$ , and  $p = k + 1$ . We would like to determine if the subset of regressor variables  $x_1, x_2, \dots, x_r$  ( $r < k$ ) contribute significantly to the regression model. Let the vector of regression coefficients be partitioned as follows:

$$\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix}$$

where  $\boldsymbol{\beta}_1$  is  $r \times 1$ , and  $\boldsymbol{\beta}_2$  is  $(p-r) \times 1$ . We wish to test the hypotheses

$$\begin{aligned} H_0: \boldsymbol{\beta}_1 &= \mathbf{0} \\ H_1: \boldsymbol{\beta}_1 &\neq \mathbf{0} \end{aligned} \quad (2.30)$$

The model may be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon} \quad (2.31)$$

where  $\mathbf{X}_1$  represents the columns of  $\mathbf{X}$  associated with  $\boldsymbol{\beta}_1$ , and  $\mathbf{X}_2$  represents the columns of  $\mathbf{X}$  associated with  $\boldsymbol{\beta}_2$ .

For the **full model** (including both  $\boldsymbol{\beta}_1$  and  $\boldsymbol{\beta}_2$ ), we know that  $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}$ . Also, the regression sum of squares for all variables including the intercept is

$$SS_R(\boldsymbol{\beta}) = \mathbf{b}'\mathbf{X}'\mathbf{y} \quad (p \text{ degrees of freedom})$$

and

$$MS_E = \frac{\mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y}}{n - p}$$

$SS_R(\boldsymbol{\beta})$  is called the regression sum of squares due to  $\boldsymbol{\beta}$ . To find the contribution of the terms in  $\boldsymbol{\beta}_1$  to the regression, fit the model assuming the null hypothesis  $H_0: \boldsymbol{\beta}_1 = \mathbf{0}$  to be true. The **reduced model** is found from Equation 2.31 with  $\boldsymbol{\beta}_1 = \mathbf{0}$ :

$$\mathbf{y} = \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon} \quad (2.32)$$

The least squares estimator of  $\boldsymbol{\beta}_2$  is  $\mathbf{b}_2 = (\mathbf{X}'_2\mathbf{X}_2)^{-1} \mathbf{X}'_2\mathbf{y}$ , and

$$SS_R(\boldsymbol{\beta}_2) = \mathbf{b}'_2\mathbf{X}'_2\mathbf{y} \quad (p - r \text{ degrees of freedom}) \quad (2.33)$$

The increase in the regression sum of squares due to  $\boldsymbol{\beta}_1$  given that  $\boldsymbol{\beta}_2$  is already in the model is

$$SS_R(\boldsymbol{\beta}_1|\boldsymbol{\beta}_2) = SS_R(\boldsymbol{\beta}) - SS_R(\boldsymbol{\beta}_2) \quad (2.34)$$

This sum of squares has  $r$  degrees of freedom. It is often called the **extra sum of squares** due to  $\boldsymbol{\beta}_1$ . Note that  $SS_R(\boldsymbol{\beta}_1|\boldsymbol{\beta}_2)$  is the increase in the regression sum of squares due to including the variables  $x_1, x_2, \dots, x_r$  in the model. Now  $SS_R(\boldsymbol{\beta}_1|\boldsymbol{\beta}_2)$  is independent of  $MS_E$ ,

and the null hypothesis  $\beta_1 = \mathbf{0}$  may be tested by the statistic

$$F_0 = \frac{SS_R(\beta_1|\beta_2)/r}{MS_E} \quad (2.35)$$

If  $F_0 > F_{\alpha, r, n-p}$ , we reject  $H_0$ , concluding that at least one of the parameters in  $\beta_1$  is not zero and, consequently, at least one of the variables  $x_1, x_2, \dots, x_r$  in  $\mathbf{X}_1$  contributes significantly to the regression model (one could also use the  $P$ -value approach). Some authors call the test in Equation 2.35 a **partial  $F$ -test**.

The partial  $F$ -test is very useful. We can use it to measure the contribution of  $x_j$  as if it were the last variable added to the model by computing

$$SS_R(\beta_j|\beta_0, \beta_1, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_k)$$

This is the increase in the regression sum of squares due to adding  $x_j$  to a model that already includes  $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k$ . Note that the partial  $F$ -test on a single variable  $x_j$  is equivalent to the  $t$ -test in Equation 2.27. However, the partial  $F$ -test is a more general procedure in that we can measure the effect of sets of variables. This procedure is used often in response surface work. For example, suppose that we are considering fitting the second-order model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \varepsilon$$

and we wish to test the contribution of the second-order terms over and above the contribution from the first-order model. Therefore, the hypotheses of interest are

$$H_0: \beta_{11} = \beta_{22} = \beta_{12} = 0$$

$$H_1: \beta_{11} \neq 0 \quad \text{and/or} \quad \beta_{22} \neq 0 \quad \text{and/or} \quad \beta_{12} \neq 0$$

In the notation of this section,  $\beta'_1 = [\beta_{11}, \beta_{22}, \beta_{12}]$  and  $\beta'_2 = [\beta_0, \beta_1, \beta_2]$ , and the columns of  $\mathbf{X}_1$  and  $\mathbf{X}_2$  are the columns of the original  $\mathbf{X}$  matrix associated with the quadratic and linear terms in the model, respectively.

**Example 2.5 Partial  $F$ -tests for the Transistor Gain Data** Consider the transistor gain data in Example 2.1. Suppose that we wish to investigate the contribution of the variable  $x_2$  (dose) to the model. That is, the hypotheses we wish to test are

$$H_0: \beta_2 = 0$$

$$H_1: \beta_2 \neq 0$$

This will require the extra sum of squares due to  $\beta_2$ , or

$$\begin{aligned} SS_R(\beta_2|\beta_1, \beta_0) &= SS_R(\beta_0, \beta_1, \beta_2) - SS_R(\beta_0, \beta_1) \\ &= SS_R(\beta_1, \beta_2|\beta_0) - SS_R(\beta_1|\beta_0) \end{aligned}$$



Now from Example 2.3 where we tested for significance of regression, we have (from Table 2.6)

$$SS_R(\beta_1, \beta_2|\beta_0) = 651,966.1$$

This sum of squares has two degrees of freedom. The reduced model is

$$y = \beta_0 + \beta_1 x_1 + \varepsilon$$

The least squares fit for this model is

$$\hat{y} = 1245.8 + 323.6x_1$$

and the regression sum of squares for this model (with one degree of freedom) is

$$SS_R(\beta_1|\beta_0) = 630,967.9$$

Therefore,

$$\begin{aligned} SS_R(\beta_2|\beta_0, \beta_1) &= 651,966.1 - 630,967.9 \\ &= 20,998.2 \end{aligned}$$

with  $2-1 = 1$  degree of freedom. This is the increase in the regression sum of squares that results from adding  $x_2$  to a model already containing  $x_1$  (or the extra sum of squares that results from adding  $x_2$  to a model containing  $x_1$ ). To test  $H_0: \beta_2 = 0$ , from the test statistic we obtain

$$F_0 = \frac{SS_R(\beta_2|\beta_0, \beta_1)/1}{MS_E} = \frac{20,998.2/1}{1220.1} = 17.21$$

Note that  $MS_E$  from the full model (Table 2.6) is used in the denominator of  $F_0$ . Now because  $F_{0.05,1,11} = 4.84$ , we reject  $H_0: \beta_2 = 0$  and conclude that  $x_2$  (dose) contributes significantly to the model.

Because this partial  $F$ -test involves only a single regressor, it is equivalent to the  $t$ -test introduced earlier, because the square of a  $t$  random variable with  $\nu$  degrees of freedom is an  $F$  random variable with 1 and  $\nu$  degrees of freedom. To see this, recall that the  $t$ -statistic for  $H_0: \beta_2 = 0$  resulted in  $t_0 = -4.15$  and that  $t_0^2 = (-4.15)^2 = 17.22 \cong F_0$ .

## 2.5 CONFIDENCE INTERVALS IN MULTIPLE REGRESSION

It is often beneficial to construct confidence interval estimates for the regression coefficients  $\{\beta_j\}$  and for other quantities of interest from the regression model. The development of a procedure for obtaining these confidence intervals requires that we assume the errors  $\{\varepsilon_i\}$  to be normally and independently distributed with mean zero and variance  $\sigma^2$ , the same assumption made in the section on hypothesis testing (Section 2.4).

### 2.5.1 Confidence Intervals on the Individual Regression Coefficients $\beta$

Because the least squares estimator  $\mathbf{b}$  is a linear combination of the observations, it follows that  $\mathbf{b}$  is normally distributed with mean vector  $\beta$  and covariance matrix  $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$ . Then each of the statistics

$$\frac{b_j - \beta_j}{\sqrt{\hat{\sigma}^2 C_{jj}}}, \quad j = 0, 1, \dots, k \quad (2.36)$$

is distributed as  $t$  with  $n - p$  degrees of freedom, where  $C_{jj}$  is the  $(j, j)$ th element of the matrix  $(\mathbf{X}'\mathbf{X})^{-1}$ , and  $\hat{\sigma}^2$  is the estimate of the error variance, obtained from Equation 2.18. Therefore, a  $100(1-\alpha)\%$  confidence interval for the regression coefficient  $\beta_j, j = 0, 1, \dots, k$ , is

$$b_j - t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2 C_{jj}} \leq \beta_j \leq b_j + t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2 C_{jj}} \quad (2.37)$$

Note that this confidence interval can also be written as

$$b_j - t_{\alpha/2, n-p} se(b_j) \leq \beta_j \leq b_j + t_{\alpha/2, n-p} se(b_j)$$

because  $se(b_j) = \sqrt{\hat{\sigma}^2 C_{jj}}$ .

**Example 2.6 The Transistor Gain Data** We will construct a 95% confidence interval for the parameter  $\beta_1$  in Example 2.1. Now  $b_1 = 323.4$ , and because  $\hat{\sigma}^2 = 1220.1$  and  $C_{11} = 0.165954$ , we find that

$$\begin{aligned} b_1 - t_{0.025, 11} \sqrt{\hat{\sigma}^2 C_{11}} &\leq \beta_1 \leq b_1 + t_{0.025, 11} \sqrt{\hat{\sigma}^2 C_{11}} \\ 323.4 - 2.201 \sqrt{(1220.1)(0.165954)} & \\ &\leq \beta_1 \leq 323.4 + 2.201 \sqrt{(1220.1)(0.165954)} \\ 323.4 - 2.201(14.2) &\leq \beta_1 \leq 323.4 + 2.201(14.2) \end{aligned}$$

and the 95% confidence interval on  $\beta_1$  is

$$292.1 \leq \beta_1 \leq 354.7$$

### 2.5.2 A Joint Confidence Region on the Regression Coefficients $\beta$

The confidence intervals in the previous subsection should be thought of as one-at-a-time intervals; that is, the confidence coefficient  $1 - \alpha$  applies only to one such interval. Some problems require that several confidence intervals be constructed from the same data. In such cases, the analyst is usually interested in specifying a confidence coefficient that applies to the entire set of confidence intervals. Such intervals are called **simultaneous confidence intervals**.

It is relatively easy to specify a **joint confidence region** for the parameters  $\beta$  in a multiple regression model. We may show that

$$\frac{(\mathbf{b} - \beta)' \mathbf{X}' \mathbf{X} (\mathbf{b} - \beta)}{pMS_E}$$

has an  $F$ -distribution with  $p$  numerator and  $n-p$  denominator degrees of freedom, and this implies that

$$P \left\{ \frac{(\mathbf{b} - \beta)' \mathbf{X}' \mathbf{X} (\mathbf{b} - \beta)}{pMS_E} \leq F_{\alpha, p, n-p} \right\} = 1 - \alpha$$

Consequently a  $100(1-\alpha)\%$  point confidence region for all the parameters in  $\beta$  is

$$\frac{(\mathbf{b} - \beta)' \mathbf{X}' \mathbf{X} (\mathbf{b} - \beta)}{pMS_E} \leq F_{\alpha, p, n-p} \quad (2.38)$$

This inequality describes an elliptically shaped region, Montgomery, Peck, and Vining (2012) and Myers (1990) demonstrate the construction of this region for  $p = 2$ . When there are only two parameters, finding this region is relatively simple; however, when more than two parameters are involved, the construction problem is considerably harder.

There are other methods for finding joint or simultaneous intervals on regression coefficients, Montgomery, Peck, and Vining (2012) and Myers (1990) discuss and illustrate many of these methods. They also present methods for finding several other types of interval estimates.

### 2.5.3 Confidence Interval on the Mean Response

We may also obtain a confidence interval on the mean response at a particular point, say,  $x_{01}, x_{02}, \dots, x_{0k}$ . Define the vector

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ x_{01} \\ x_{02} \\ \vdots \\ x_{0k} \end{bmatrix}$$

The mean response at this point is

$$\mu_{y|\mathbf{x}_0} = \beta_0 + \beta_1 x_{01} + \beta_2 x_{02} + \dots + \beta_k x_{0k} = \mathbf{x}'_0 \beta$$

The estimated mean response at this point is

$$\hat{y}(\mathbf{x}_0) = \mathbf{x}'_0 \mathbf{b} \quad (2.39)$$

This estimator is unbiased, because  $E[\hat{y}_i(\mathbf{x}_0)] = E(\mathbf{x}'_0 \mathbf{b}) = \mathbf{x}'_0 \beta = \mu_{y|\mathbf{x}_0}$ , and the variance of  $\hat{y}_i(\mathbf{x}_0)$  is

$$\text{Var}[\hat{y}(\mathbf{x}_0)] = \sigma^2 \mathbf{x}'_0 (\mathbf{X}' \mathbf{X})^{-1} \mathbf{x}_0 \quad (2.40)$$

Therefore, a  $100(1-\alpha)\%$  confidence interval on the mean response at the point  $x_{01}$ ,  $x_{02}, \dots, x_{0k}$  is

$$\begin{aligned} \hat{y}(\mathbf{x}_0) - t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2 \mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0} \\ \leq \mu_{y|\mathbf{x}_0} \leq \hat{y}(\mathbf{x}_0) + t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2 \mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0} \end{aligned} \quad (2.41)$$

**Example 2.7 The Transistor Gain Data** Suppose that we wish to find a 95% confidence interval on the mean response for the transistor gain problem for the point  $\xi_1 = 225$  min and  $\xi_2 = 4.36 \times 10^{14}$  ions. In terms of the coded variables  $x_1$  and  $x_2$ , this point corresponds to

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

because  $x_{01} = 0$  and  $x_{02} = 0$ . The estimate of the mean response at this point is computed from Equation 2.39 as

$$\hat{y}(\mathbf{x}_0) = \mathbf{x}'_0 \mathbf{b} = [1, 0, 0] \begin{bmatrix} 1242.3 \\ 323.4 \\ -54.8 \end{bmatrix} = 1242.3$$

Now from Equation 2.40, we find  $\text{Var}[\hat{y}(\mathbf{x}_0)]$  as

$$\begin{aligned} \text{Var}[\hat{y}(\mathbf{x}_0)] &= \sigma^2 \mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0 \\ &= \sigma^2 [1, 0, 0] \begin{bmatrix} 0.072027 & -0.00195 & 0.009067 \\ -0.00195 & 0.165954 & 0.000408 \\ 0.009067 & 0.000408 & 0.142876 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ &= \sigma^2 (0.072027) \end{aligned}$$

Using  $\hat{\sigma}^2 = MS_E = 1220.1$  and Equation 2.41, we find the confidence interval as

$$\begin{aligned} \hat{y}(\mathbf{x}_0) - t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2 \mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0} \leq \mu_{y|\mathbf{x}_0} \leq \hat{y}(\mathbf{x}_0) + t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2 \mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0} \\ 1242.3 - 2.201 \sqrt{1220.1(0.072027)} \leq \mu_{y|\mathbf{x}_0} \leq 1242.3 + 2.201 \sqrt{1220.1(0.072027)} \\ 1242.3 - 20.6 \leq \mu_{y|\mathbf{x}_0} \leq 1242.3 + 20.6 \end{aligned}$$

or

$$1221.7 \leq \mu_{y|\mathbf{x}_0} \leq 1262.9$$

### 2.6 PREDICTION OF NEW RESPONSE OBSERVATIONS

A regression model can be used to predict future observations on the response  $y$  corresponding to particular values of the regressor variables, say  $x_{01}, x_{02}, \dots, x_{0k}$ . If  $\mathbf{x}'_0 = [1, x_{01}, x_{02}, \dots, x_{0k}]$ , then a point estimate for the future observation  $y_0$  at the point  $x_{01}, x_{02}, \dots, x_{0k}$  is computed from Equation 2.39:

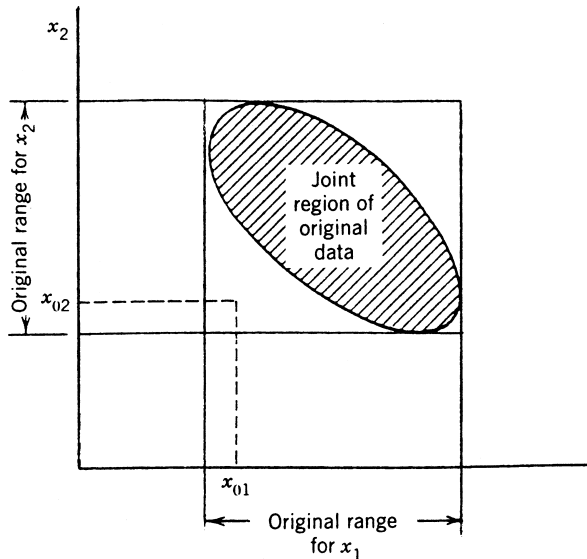
$$\hat{y}(\mathbf{x}_0) = \mathbf{x}'_0 \mathbf{b}$$

A  $100(1-\alpha)\%$  **prediction interval** for this future observation is

$$\begin{aligned} & \hat{y}(\mathbf{x}_0) - t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2(1 + \mathbf{x}'_0(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0)} \\ & \leq y_0 \leq \hat{y}(\mathbf{x}_0) + t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2(1 + \mathbf{x}'_0(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0)} \end{aligned} \quad (2.42)$$

The additional  $\hat{\sigma}^2$  term under the square root sign is due to variability of observations around the predicted mean at that location.

In predicting new observations and in estimating the mean response at a given point  $x_{01}, x_{02}, \dots, x_{0k}$ , one must be careful about extrapolating beyond the region containing the original observations. It is very possible that a model that fits well in the region of the original data will no longer fit well outside of that region. In multiple regression it is often easy to inadvertently extrapolate, since the levels of the variables  $(x_{i1}, x_{i2}, \dots, x_{ik}), i = 1, 2, \dots, n$ , jointly define the region containing the data. As an example, consider Fig. 2.3, which illustrates the region containing the observations for a two-variable regression model.



**Figure 2.3** An example of extrapolation in multiple regression.

Note that the point  $(x_{01}, x_{02})$  lies within the ranges of both regressor variables  $x_1$  and  $x_2$ , but it is outside the region of the original observations. Thus, either predicting the value of a new observation or estimating the mean response at this point is an extrapolation of the original regression model.

**Example 2.8 A Prediction Interval for the Transistor Gain Data** Suppose that we wish to find a 95% prediction interval on the next observation on the transistor gain at the point  $\xi_1 = 225$  min and  $\xi_2 = 4.36 \times 10^{14}$  ions. In the coded variables, this point is  $x_{01} = 0$  and  $x_{02} = 0$ , so that, as in Example 2.7,  $\mathbf{x}'_0 = [1, 0, 0]$ , and the predicted value of the gain at this point is  $\hat{y}_i(\mathbf{x}_0) = \mathbf{x}'_0 \mathbf{b} = 1242.3$ . From Example 2.7 we know that

$$\mathbf{x}'_0(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0 = 0.072027$$

Therefore, using Equation 2.42 we can find the 95% prediction interval on  $y_0$  as follows:

$$\begin{aligned} \hat{y}(\mathbf{x}_0) - t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2(1 + \mathbf{x}'_0(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0)} &\leq y_0 \leq \hat{y}(\mathbf{x}_0) + t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2(1 + \mathbf{x}'_0(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0)} \\ 1242.3 - 2.201 \sqrt{1220.1(1 + 0.072027)} &\leq y_0 \leq 1242.3 + 2.201 \sqrt{1220.1(1 + 0.072027)} \\ 1242.3 - 79.6 &\leq y_0 \leq 1242.3 + 79.6 \end{aligned}$$

or

$$1162.7 \leq y_0 \leq 1321.9$$

If we compare the width of the prediction interval at this point with the width of the confidence interval on the mean gain at the same point from Example 2.7, we observe that the prediction interval is much wider. This reflects the fact that there is more uncertainty in our prediction of an individual future value of a random variable than to estimate the mean of the probability distribution from which that future observation will be drawn.

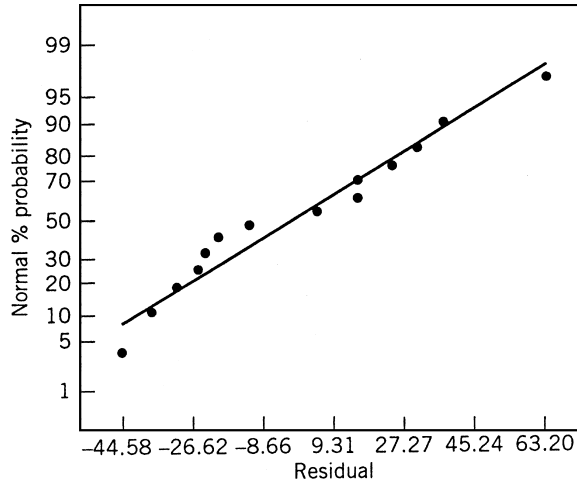
## 2.7 MODEL ADEQUACY CHECKING

It is always necessary to (a) examine the fitted model to ensure that it provides an adequate approximation to the true system and (b) verify that none of the least squares regression assumptions are violated. Proceeding with exploration and optimization of a fitted response surface will likely give poor or misleading results unless the model provides an adequate fit. In this section we present several techniques for checking model adequacy.

### 2.7.1 Residual Analysis

The residuals from the least squares fit, defined by  $e_i = y_i - \hat{y}_i$ ,  $i = 1, 2, \dots, n$ , play an important role in judging model adequacy. The residuals from the transistor gain regression model of Example 2.1 are shown in column 3 of Table 2.3.

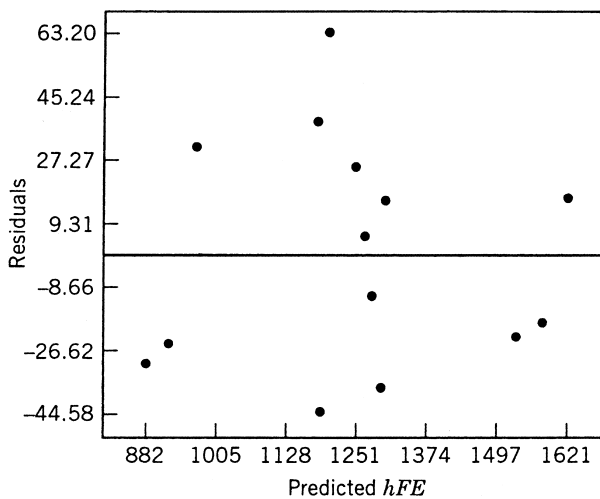
A check of the normality assumption may be made by constructing a normal probability plot of the residuals, as in Fig. 2.4. If the residuals plot approximately along a straight



**Figure 2.4** Normal probability plot of residuals, Example 2.1.

line, then the normality assumption is satisfied. Figure 2.4 reveals no apparent problem with normality. The straight line in this normal probability plot was determined by eye, concentrating on the central portion of the data. When this plot indicates problems with the normality assumption, we often transform the response variable as a remedial measure. For more details, see Montgomery, Peck, and Vining (2012) and Myers (1990).

Figure 2.5 presents a plot of residuals  $e_i$  versus the predicted response  $\hat{y}_i$  (this plot is also shown in the JMP output in Fig. 2.2). The general impression is that the residuals scatter randomly on the display, suggesting that the variance of the original observations is constant for all values of  $y$ . If the variance of the response depends on the mean level of  $y$ , then this plot will often exhibit a funnel-shaped pattern. This is also suggestive of the need for transformation of the response variable  $y$ .



**Figure 2.5** Plot of residuals versus predicted response  $\hat{y}_i$ , Example 2.1.

It is also useful to plot the residuals in time or run order and versus each of the individual regressors. Nonrandom patterns on these plots would indicate model inadequacy. In some cases, transformations may stabilize the situation. See Montgomery, Peck, and Vining (2012) and Myers (1990) and Section 2.9 for more details.

### 2.7.2 Scaling Residuals

**Standardized and Studentized Residuals** Many response surface analysts prefer to work with **scaled residuals**, as these scaled residuals often convey more information than do the ordinary least squares residuals.

One type of scaled residual is the **standardized residual**:

$$d_i = \frac{e_i}{\hat{\sigma}}, \quad i = 1, 2, \dots, n \quad (2.43)$$

where we generally use  $\hat{\sigma} = \sqrt{MS_E}$  in the computation. These standardized residuals have mean zero and approximately unit variance; consequently, they are useful in looking for **outliers**. Most of the standardized residuals should lie in the interval  $-3 \leq d_i \leq 3$ , and any observation with a standardized residual outside of this interval is potentially unusual with respect to its observed response. These outliers should be carefully examined, because they may represent something as simple as a data recording error or something of more serious concern, such as a region of the regressor variable space where the fitted model is a poor approximation to the true response surface.

The standardizing process in Equation 2.43 scales the residuals by dividing them by their average standard deviation. In some data sets, residuals may have standard deviations that differ greatly. We now present a scaling that takes this into account.

The vector of fitted values  $\hat{y}_i$  corresponding to the observed values  $y_i$  is

$$\begin{aligned} \hat{\mathbf{y}} &= \mathbf{X}\mathbf{b} \\ &= \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ &= \mathbf{H}\mathbf{y} \end{aligned} \quad (2.44)$$

The  $n \times n$  matrix  $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$  is usually called the **hat** matrix because it maps the vector of observed values into a vector of fitted values. The hat matrix and its properties play a central role in regression analysis.

The residuals from the fitted model may be conveniently written in matrix notation as

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} \quad (2.45)$$

There are several other ways to express the vector of residuals  $\mathbf{e}$  that will prove useful, including

$$\begin{aligned} \mathbf{e} &= \mathbf{y} - \mathbf{X}\mathbf{b} \\ &= \mathbf{y} - \mathbf{H}\mathbf{y} \\ &= (\mathbf{I} - \mathbf{H})\mathbf{y} \end{aligned} \quad (2.46)$$



The hat matrix has several useful properties. It is symmetric ( $\mathbf{H}' = \mathbf{H}$ ) and idempotent ( $\mathbf{H}\mathbf{H} = \mathbf{H}$ ). Similarly the matrix  $\mathbf{I} - \mathbf{H}$  is symmetric and idempotent.

The covariance matrix of the residuals is

$$\begin{aligned}\text{Var}(\mathbf{e}) &= \text{Var}[(\mathbf{I} - \mathbf{H})\mathbf{y}] \\ &= (\mathbf{I} - \mathbf{H})\text{Var}(\mathbf{y})(\mathbf{I} - \mathbf{H})' \\ &= \sigma^2(\mathbf{I} - \mathbf{H})\end{aligned}\quad (2.47)$$

because  $\text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}$  and  $\mathbf{I} - \mathbf{H}$  is symmetric and idempotent. The matrix  $\mathbf{I} - \mathbf{H}$  is generally not diagonal, so the residuals are correlated and have different variances.

The variance of the  $i$ th residual is

$$\text{Var}(e_i) = \sigma^2(1 - h_{ii}) \quad (2.48)$$

where  $h_{ii}$  is the  $i$ th diagonal element of  $\mathbf{H}$ . Because  $0 \leq h_{ii} \leq 1$ , using the residual mean square  $MS_E$  to estimate the variance of the residuals actually overestimates  $\text{Var}(e_i)$ . Furthermore, because  $h_{ii}$  is a measure of the location of the  $i$ th point in  $x$ -space, the variance of  $e_i$  depends upon where the point  $\mathbf{x}_i$  lies. Generally, residuals near the center of the  $x$ -space have larger variance than do residuals at more remote locations. Violations of model assumptions are more likely at remote points, and these violations may be hard to detect from inspection of  $e_i$  (or  $d_i$ ) because their residuals will usually be smaller.

We recommend taking this inequality of variance into account when scaling the residuals. We suggest plotting the **studentized residuals**:

$$r_i = \frac{e_i}{\sqrt{\hat{\sigma}^2(1 - h_{ii})}}, \quad i = 1, 2, \dots, n \quad (2.49)$$

with  $\hat{\sigma}^2 = MS_E$  instead of  $e_i$  (or  $d_i$ ). The studentized residuals have constant variance  $\text{Var}(r_i) = 1$  regardless of the location of  $\mathbf{x}_i$  when the form of the model is correct. In many situations the variance of the residuals stabilizes, particularly for large data sets. In these cases there may be little difference between the standardized and studentized residuals. Thus standardized and studentized residuals often convey equivalent information. However, because any point with a large residual and a large  $h_{ii}$  is potentially highly influential on the least squares fit, examination of the studentized residuals is generally recommended.

**PRESS Residuals** The **prediction error sum of squares** (PRESS) proposed by Allen (1971, 1974) provides a useful residual scaling. To calculate PRESS, select an observation—for example,  $i$ . Fit the regression model to the remaining  $n-1$  observations and use this equation to predict the withheld observation  $y_i$ . Denoting this predicted value by  $\hat{y}_{(i)}$ , we may find the prediction error for point  $i$  as  $e_{(i)} = y_i - \hat{y}_{(i)}$ . The prediction error is often called the  $i$ th PRESS residual. This procedure is repeated for each observation  $i = 1, 2, \dots, n$ , producing a set of  $n$  PRESS residuals  $e_{(1)}, e_{(2)}, \dots, e_{(n)}$ . Then the PRESS statistic is defined as the sum of squares of the  $n$  PRESS residuals as in

$$\text{PRESS} = \sum_{i=1}^n e_{(i)}^2 = \sum_{i=1}^n [y_i - \hat{y}_{(i)}]^2 \quad (2.50)$$

Thus PRESS uses each possible subset of  $n-1$  observations as an estimation data set, and every observation in turn is used to form a prediction data set.

It would initially seem that calculating PRESS requires fitting  $n$  different regressions. However, it is possible to calculate PRESS from the results of a single least squares fit to all  $n$  observations. It turns out that the  $i$ th PRESS residual is

$$e_{(i)} = \frac{e_i}{1 - h_{ii}} \quad (2.51)$$

Thus, because PRESS is just the sum of the squares of the PRESS residuals, a simple computing formula is

$$\text{PRESS} = \sum_{i=1}^n \left( \frac{e_i}{1 - h_{ii}} \right)^2 \quad (2.52)$$

From Equation 2.51 it is easy to see that the PRESS residual is just the ordinary residual weighted according to the diagonal elements of the hat matrix  $h_{ii}$ . Data points for which  $h_{ii}$  are large will have large PRESS residuals. These observations will generally be **high influence** points. Generally, a large difference between the ordinary residual and the PRESS residual will indicate a point where the model fits the data well, but a model built without that point predicts poorly. In the next section we will discuss some other measures of influence.

The variance of the  $i$ th PRESS residual is

$$\begin{aligned} \text{Var}[e_{(i)}] &= \text{Var} \left[ \frac{e_i}{1 - h_{ii}} \right] \\ &= \frac{1}{(1 - h_{ii})^2} \sigma^2 (1 - h_{ii}) \\ &= \frac{\sigma^2}{1 - h_{ii}} \end{aligned} \quad (2.53)$$

so that standardized PRESS residual is

$$\begin{aligned} \frac{e_{(i)}}{\sqrt{\text{Var}[e_{(i)}]}} &= \frac{e_i/(1 - h_{ii})}{\sqrt{\sigma^2/(1 - h_{ii})}} \\ &= \frac{e_i}{\sqrt{\sigma^2(1 - h_{ii})}} \end{aligned}$$

which, if we use  $MS_E$  to estimate  $\sigma^2$ , is just the studentized residual discussed previously.

Finally, we note that PRESS can be used to compute an approximate  $R^2$  for prediction, say

$$R_{\text{prediction}}^2 = 1 - \frac{\text{PRESS}}{SS_T} \quad (2.54)$$

This statistic gives some indication of the predictive capability of the regression model. For the transistor gain model we can compute the PRESS residuals using the ordinary residuals

and the values of  $h_{ii}$  found in Table 2.3. The corresponding value of the PRESS statistic is  $\text{PRESS} = 22,225.0$ . Then

$$\begin{aligned} R^2_{\text{prediction}} &= 1 - \frac{\text{PRESS}}{SS_T} \\ &= 1 - \frac{22,225.0}{665,387.2} \\ &= 0.9666 \end{aligned}$$

Therefore we could expect this model to “explain” about 96.66% of the variability in predicting new observations, as compared to the approximately 97.98% of the variability in the original data explained by the least squares fit. The overall predictive capability of the model based on this criterion seems very satisfactory.

**R Student** The studentized residual  $r_i$  discussed above is often considered an outlier diagnostic. It is customary to use  $MS_E$  as an estimate of  $\sigma^2$  in computing  $r_i$ . This is referred to as **internal scaling** of the residual, because  $MS_E$  is an internally generated estimate of  $\sigma^2$  obtained from fitting the model to all  $n$  observations. Another approach would be to use an estimate of  $\sigma^2$  based on a data set with the  $i$ th observation removed. Denote the estimate of  $\sigma^2$  so obtained by  $S^2_{(i)}$ . We can show that

$$S^2_{(i)} = \frac{(n-p)MS_E - e_i^2/(1-h_{ii})}{n-p-1} \quad (2.55)$$

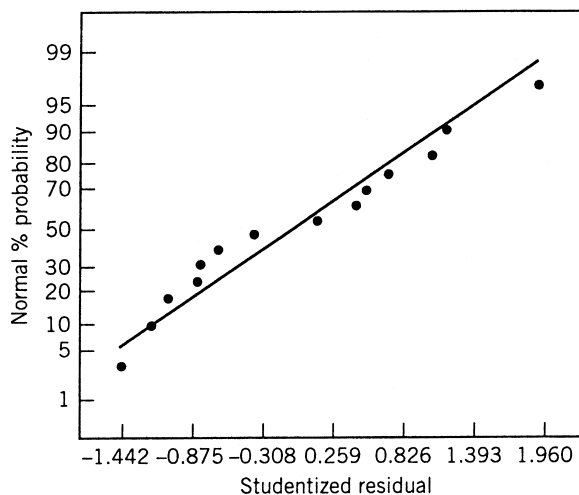
The estimate of  $\sigma^2$  in Equation 2.55 is used instead of  $MS_E$  to produce an **externally studentized residual**, usually called **R-student**, given by

$$t_i = \frac{e_i}{\sqrt{S^2_{(i)}(1-h_{ii})}}, \quad i = 1, 2, \dots, n \quad (2.56)$$

Some computer packages refer to *R*-student as the outlier *T*.

In many situations,  $t_i$  will differ little from the studentized residual  $r_i$ . However, if the  $i$ th observation is influential, then  $S^2_{(i)}$  can differ significantly from  $MS_E$ , and thus *R*-student will be more sensitive to this point. Furthermore, under the standard assumptions,  $t_i$  has a  $t_{n-p-1}$ -distribution. Thus *R*-student offers a more formal procedure for outlier detection via hypothesis testing. One could use a simultaneous inference procedure called the **Bonferroni approach** and compare all  $n$  values of  $[t_i]$  with  $t_{(\alpha/2n), n-p-1}$  to provide guidance regarding outliers. However, it is our view that a formal approach is usually not necessary and that only relatively crude cutoff values need be considered. In general, a diagnostic view as opposed to a strict statistical hypothesis-testing view is best. Furthermore, detection of outliers needs to be considered simultaneously with detection of influential observations.

**Example 2.9 The Transistor Gain Data** Table 2.3 presents the studentized residuals  $r_i$  and the *R*-student values  $t_i$  defined in Equations 2.50 and 2.56 for the transistor gain regression model. None of these values are large enough to cause any concern regarding outliers.



**Figure 2.6** Normal probability plot of the studentized residuals for the transistor gain data.

Figure 2.6 is a normal probability plot of the studentized residuals. It conveys exactly the same information as the normal probability plot of the ordinary residuals  $e_i$  in Fig. 2.4. This is because most of the  $h_{ii}$ -values are similar and there are no unusually large residuals. In some applications, however, the  $h_{ii}$  can differ considerably, and in those cases plotting the studentized residuals is the best approach.

### 2.7.3 Influence Diagnostics

We occasionally find that a small subset of the data exerts a disproportionate influence on the fitted regression model. That is, parameter estimates or predictions may depend more on the influential subset than on the majority of the data. We would like to locate these influential points and assess their impact on the model. If these influential points are “bad” values, then they should be eliminated. On the other hand, there may be nothing wrong with these points, but if they control key model properties, we would like to know it, because it could affect the use of the model. In this subsection we describe and illustrate several useful measure of influence.

**Leverage Points** The disposition of points in  $x$ -space is important in determining model properties. In particular remote observations potentially have disproportionate leverage on the parameter estimates, the predicted values, and the usual summary statistics.

The hat matrix  $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$  is very useful in identifying influential observations. As noted earlier,  $\mathbf{H}$  determines the variances and covariances of  $\hat{\mathbf{y}}$  and  $\mathbf{e}$ , because  $\text{Var}(\hat{\mathbf{y}}) = \sigma^2 \mathbf{H}$  and  $\text{Var}(\mathbf{e}) = \sigma^2(\mathbf{I} - \mathbf{H})$ . The elements  $h_{ii}$  of  $\mathbf{H}$  may be interpreted as the amount of leverage exerted by  $y_j$  on  $\hat{y}_i$ . Thus inspection of the elements of  $\mathbf{H}$  can reveal points that are potentially influential by virtue of their location in  $x$ -space. Attention is usually focused on the diagonal elements  $h_{ii}$ . Because  $\sum_{i=1}^n h_{ii} = \text{rank}(\mathbf{H}) = \text{rank}(\mathbf{X}) = p$ , the average size of the diagonal element of the matrix  $\mathbf{H}$  is  $p/n$ . As a rough guideline, then, if a diagonal element  $h_{ii}$  is greater than  $2p/n$ , observation  $i$  is a high-leverage point. To apply this to the transistor gain data in Example 2.1, note that  $2p/n = 2(3)/14 = 0.43$ . Table 2.3 gives the hat

diagonals  $h_{ii}$  for the first-order model; and because none of the  $h_{ii}$  exceed 0.43, we conclude that there are no leverage points in these data. Further properties and uses of the elements of the hat matrix in regression diagnostics are discussed by Belsley, Kah, and Welsch (1980).

**Influence on Regression Coefficients** The hat diagonals will identify points that are potentially influential due to their location in  $x$ -space. It is desirable to consider both the location of the point and the response variable in measuring influence. Cook (1977, 1979) has suggested using a measure of the squared distance between the least squares estimate based on all  $n$  points  $\mathbf{b}$  and the estimate obtained by deleting the  $i$ th point, say  $\mathbf{b}_{(i)}$ . This distance measure can be expressed in a general form as

$$D_i(\mathbf{M}, c) = \frac{(\mathbf{b}_{(i)} - \mathbf{b})' \mathbf{M} (\mathbf{b}_{(i)} - \mathbf{b})}{c}, \quad i = 1, 2, \dots, n \quad (2.57)$$

The usual choices of  $\mathbf{M}$  and  $c$  are  $\mathbf{M} = \mathbf{X}'\mathbf{X}$  and  $c = pMS_E$ , so that Equation 2.57 becomes

$$D_i(\mathbf{M}, c) \equiv D_i = \frac{(\mathbf{b}_{(i)} - \mathbf{b})' \mathbf{X}'\mathbf{X} (\mathbf{b}_{(i)} - \mathbf{b})}{pMS_E}, \quad i = 1, 2, \dots, n \quad (2.58)$$

Points with large values of  $D_i$  have considerable influence on the least squares estimates  $\mathbf{b}$ . The magnitude of  $D_i$  may be assessed by comparing it with  $F_{\alpha, p, n-p}$ . If  $D_i \simeq F_{0.5, p, n-p}$ , then deleting point  $i$  would move  $\mathbf{b}$  to the boundary of a 50% confidence region for  $\boldsymbol{\beta}$  based on the complete data set. This is a large displacement and indicates that the least squares estimate is sensitive to the  $i$ th data point. Because  $F_{0.5, p, n-p} \simeq 1$ , we usually consider points for which  $D_i > 1$  to be influential. Practical experience has shown the cutoff value of 1 works well in identifying influential points.

The statistic  $D_i$  may be rewritten as

$$D_i = \frac{r_i^2}{p} \frac{\text{Var}[\hat{y}(\mathbf{x}_i)]}{\text{Var}(e_i)} = \frac{r_i^2}{p} \frac{h_{ii}}{1 - h_{ii}}, \quad i = 1, 2, \dots, n \quad (2.59)$$

Thus we see that, apart from the constant  $p$ ,  $D_i$  is the product of the square of the  $i$ th studentized residual and  $h_{ii}/(1-h_{ii})$ . This ratio can be shown to be the distance from the vector  $\mathbf{x}_i$  to the centroid of the remaining data. Thus  $D_i$  is made up of a component that reflects how well the model fits the  $i$ th observation  $y_i$  and a component that measures how far that point is from the rest of the data. Either component (or both) may contribute to a larger value of  $D_i$ .

Table 2.3 presents the values of  $D_i$  for the first-order model fit to the transistor gain data in Example 2.1. None of these values of  $D_i$  exceed 1, so there is no strong evidence of influential observations in these data.

### 2.7.4 Testing for Lack of Fit

In RSM, usually we are fitting the regression model to data from a designed experiment. It is frequently useful to obtain two or more observations (replicates) on the response at the

---

\*The distance measure  $D_i$  is not an  $F$  random variable, but is compared with an  $F$ -value because of the similarity of  $D_i$  to the normal theory confidence ellipsoid (Eq. 2.38).

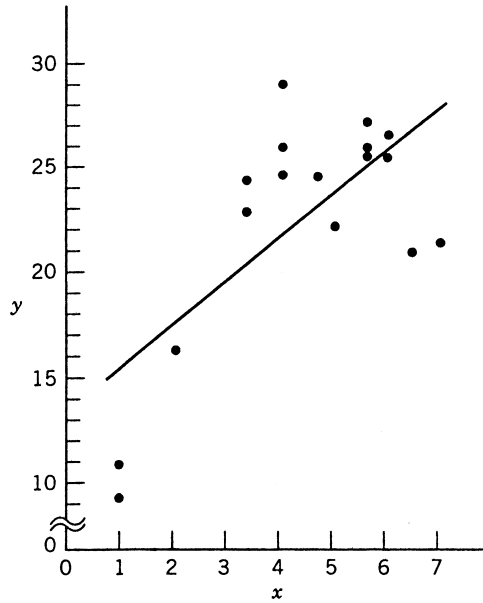


Figure 2.7 Lack of fit of the straight-line model.

same settings of the independent or regressor variables. When this has been done, we may conduct a formal test for the lack of fit on the regression model. For example, consider the data in Fig. 2.7. There is some indication that the straight-line fit is not very satisfactory, and it would be helpful to have a statistical test to determine if there is systematic curvature present.

The lack-of-fit test requires that we have **true replicates** on the response  $y$  for at least one set of levels on the regressors  $x_1, x_2, \dots, x_k$ . These are not just duplicate readings or measurements of  $y$ . For example, suppose that  $y$  is product viscosity and there is only one regressor  $x$  (temperature). True replication consists of running  $n_i$  separate experiments (usually in random order) at  $x = x_i$  and observing viscosity, not just running a single experiment at  $x_i$  and measuring viscosity  $n_i$  times. The readings obtained from the latter procedure provide information mostly on the variability of the method of measuring viscosity. The error variance  $\sigma^2$  includes measurement error, variability in the process over time, and variability associated with reaching and maintaining the same temperature level in different experiments. These replicate points are used to obtain a model-independent estimate of  $\sigma^2$ .

Suppose that we have  $n_i$  observations on the response at the  $i$ th level of the regressors  $\mathbf{x}_i$ ,  $i = 1, 2, \dots, m$ . Let  $y_{ij}$  denote the  $j$ th observation on the response at  $\mathbf{x}_i$ ,  $i = 1, 2, \dots, m$ , and  $j = 1, 2, \dots, n_i$ . There are  $n = \sum_{i=1}^m n_i$  observations altogether. The test procedure involves partitioning the residual sum of squares into two components, say

$$SS_E = SS_{PE} + SS_{LOF}$$

where  $SS_{PE}$  is the sum of squares due to pure error and  $SS_{LOF}$  is the sum of squares due to lack of fit.

To develop this partitioning of  $SS_E$ , note that the  $(i, j)$ th residual is

$$y_{ij} - \hat{y}_i = (y_{ij} - \bar{y}_i) + (\bar{y}_i - \hat{y}_i) \quad (2.60)$$

where  $\bar{y}_i$  is the average of the  $n_i$  observations at  $\mathbf{x}_i$ . Squaring both sides of Equation 2.60 and summing over  $i$  and  $j$  yields

$$\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{y}_i)^2 = \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 + \sum_{i=1}^m n_i (\bar{y}_i - \hat{y}_i)^2 \quad (2.61)$$

The left-hand side of Equation 2.61 is the usual residual sum of squares. The two components on the right-hand side measure pure error and lack of fit. We see that the pure error sum of squares

$$SS_{PE} = \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 \quad (2.62)$$

is obtained by computing the corrected sum of squares of the repeat observations at each level of  $\mathbf{x}$  and then pooling over the  $m$  levels of  $\mathbf{x}$ . If the assumption of constant variance is satisfied, this is a **model-independent** measure of pure error, because only the variability of the  $y$ s at each  $\mathbf{x}_i$  level is used to compute  $SS_{PE}$ . Because there are  $n_i - 1$  degrees of freedom for pure error at each level  $\mathbf{x}_i$ , the total number of degrees of freedom associated with the pure error sum of squares is

$$\sum_{i=1}^m (n_i - 1) = n - m \quad (2.63)$$

The sum of squares for lack of fit,

$$SS_{LOF} = \sum_{i=1}^m n_i (\bar{y}_i - \hat{y}_i)^2 \quad (2.64)$$

is a weighted sum of squared deviations between the mean response  $\bar{y}_i$  at each  $\mathbf{x}_i$  level and the corresponding fitted value. If the fitted values  $\hat{y}_i$  are close to the corresponding average responses  $\bar{y}_i$ , then there is a strong indication that the regression function is linear. If the  $\hat{y}_i$  deviate greatly from the  $\bar{y}_i$ , then it is likely that the regression function is not linear. There are  $m - p$  degrees of freedom associated with  $SS_{LOF}$ , because there are  $m$  levels of  $\mathbf{x}$  but  $p$  degrees of freedom are lost because  $p$  parameters must be estimated for the model. Computationally we usually obtain  $SS_{LOF}$  by subtracting  $SS_{PE}$  from  $SS_E$ .

The test statistic for lack of fit is

$$F_0 = \frac{SS_{LOF}/(m - p)}{SS_{PE}/(n - m)} = \frac{MS_{LOF}}{MS_{PE}} \quad (2.65)$$

The expected value of  $MS_{PE}$  is  $\sigma^2$ , and the expected value of  $MS_{LOF}$  is

$$E(MS_{LOF}) = \sigma^2 + \frac{\sum_{i=1}^m n_i \left[ E(y_i) - \beta_0 - \sum_{j=1}^k \beta_j x_{ij} \right]^2}{m-2} \quad (2.66)$$

If the true regression function is linear, then  $E(y_i) = \beta_0 + \sum_{j=1}^k \beta_j x_{ij}$ , and the second term of Equation 2.66 is zero, resulting in  $E(MS_{LOF}) = \sigma^2$ . However, if the true regression function is not linear, then  $E(y_i) \neq \beta_0 + \sum_{j=1}^k \beta_j x_{ij}$ , and  $E(MS_{LOF}) > \sigma^2$ . Furthermore, if the true regression function is linear, then the statistic  $F_0$  follows the  $F_{m-p, n-m}$ -distribution. Therefore, to test for lack of fit, we would compute the test statistic  $F_0$  and conclude that the regression function is not linear if  $F_0 > F_{\alpha, m-p, n-m}$ .

This test procedure may be easily introduced into the analysis of variance conducted for significance of regression. If we conclude that the regression function is not linear, then the tentative model must be abandoned and attempts made to find a more appropriate equation. Alternatively, if  $F_0$  does not exceed  $F_{\alpha, m-p, n-m}$ , there is no strong evidence of lack of fit, and  $MS_{PE}$  and  $MS_{LOF}$  are often combined to estimate  $\sigma^2$ .

Ideally, we find that the  $F$ -ratio for lack of fit is not significant and the hypothesis of significance of regression is rejected. Unfortunately, this does not guarantee that the model will be satisfactory as a prediction equation. Unless the variation of the predicted values is large relative to the random error, the model is not estimated with sufficient precision to yield satisfactory predictions. That is, the model may have been fitted to the errors only. Some analytical work has been done on developing criteria for judging the adequacy of the regression model from a prediction point of view. See Box and Wetz (1973), Ellerton (1978), Gunst and Mason (1979), Hill, Judge, and Fomby (1978), and Suich and Derringer (1977). Box and Wetz's work suggests that the observed  $F$ -ratio must be at least four or five times the critical value from the  $F$ -table if the regression model is to be useful as a predictor—that is, if the spread of predicted values is to be large relative to the noise.

A relatively simple measure of potential prediction performance is found by comparing the range of the fitted values  $\hat{y}$  (i.e.,  $\hat{y}_{\max} - \hat{y}_{\min}$ ) with their average standard error. It can be shown that, regardless of the form of the model, the average variance of the fitted values is

$$\overline{\text{Var}(\hat{y})} = \frac{1}{n} \sum_{i=1}^n \text{Var}[\hat{y}(\mathbf{x}_i)] = \frac{p\sigma^2}{n} \quad (2.67)$$

where  $p$  is the number of parameters in the model. In general, the model is not likely to be a satisfactory predictor unless the range of the fitted values  $\hat{y}_i$  is large relative to their average estimated standard error  $\sqrt{(p\hat{\sigma}^2)/n}$ , where  $\hat{\sigma}^2$  is a model-independent estimate of the error variance.



**TABLE 2.7 Analysis of Variance for Example 2.10**

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	$F_0$	$P$ -Value
Regression	234.7087	1	234.7087		
Residual	252.9039	15	16.8603		
Lack of fit	237.3407	8	29.6676	13.34	0.0013
Pure error	15.5632	7	2.2233		
Total	487.6126	16			

**Example 2.10 Testing Lack of Fit** The data in Fig. 2.7 are shown below:

$x$	1.0	1.0	2.0	3.3	3.3	4.0	4.0	4.0	4.7	5.0
$y$	10.84	9.30	16.35	22.88	24.35	24.56	25.86	29.46	24.59	22.25
$x$	5.6	5.6	5.6	6.0	6.0	6.5	6.92			
$y$	25.90	27.20	25.61	25.45	26.56	21.03	21.46			

The straight-line fit is  $\hat{y} = 13.301 + 2.108x$ , with  $SS_T = 487.6126$ ,  $SS_R = 234.7087$ , and  $SS_E = 252.9039$ . Note that there are 10 distinct levels of  $x$ , with repeat points at  $x = 1.0$ ,  $x = 3.3$ ,  $x = 4.0$ ,  $x = 5.6$ , and  $x = 6.0$ . The pure sum of squares is computed using the repeat points as follows:

Level of $x$	$\sum_j (y_{ij} - \bar{y}_i)^2$	Degrees of Freedom
1.0	1.1858	1
3.3	1.0805	1
4.0	11.2467	2
5.6	1.4341	2
6.0	0.6161	1
Total	15.5632	7

The lack-of-fit sum of squares is found by subtraction as

$$\begin{aligned} SS_{LOF} &= SS_E - SS_{PE} \\ &= 252.9039 - 15.5632 = 237.3407 \end{aligned}$$

with  $m - p = 10 - 2 = 8$  degrees of freedom. The analysis of variance incorporating the lack-of-fit test is shown in Table 2.7. The lack-of-fit test statistic is  $F_0 = 13.34$ , and because the  $P$ -value for this test statistic is very small, we reject the hypothesis that the tentative model adequately describes the data.

## 2.8 FITTING A SECOND-ORDER MODEL

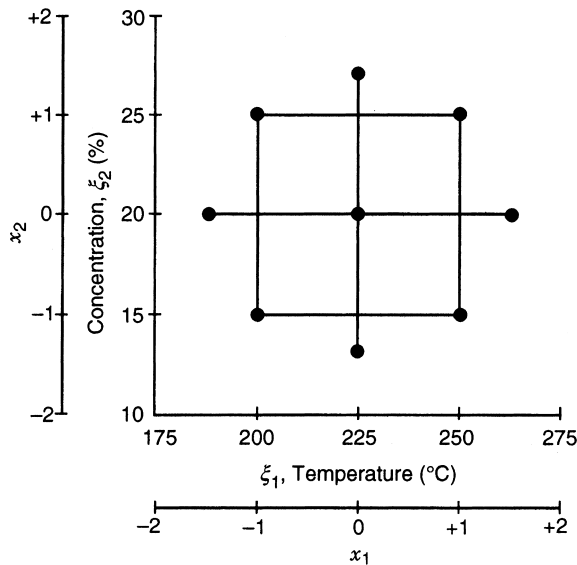
Many applications of response surface methodology involve fitting and checking the adequacy of a second-order model. In this section we present a complete example of this process.

**TABLE 2.8 Central Composite Design for the Chemical Process Example**

Observation	Run	A		B		
		Temperature (°C)	Conc. (%)	$x_1$	$x_2$	$y$
		$\xi_1$	$\xi_2$			
1	4	200	15	-1	-1	43
2	12	250	15	1	-1	78
3	11	200	25	-1	1	69
4	5	250	25	1	1	73
5	6	189.65	20	-1.414	0	48
6	7	260.35	20	1.414	0	78
7	1	225	12.93	0	-1.414	65
8	3	225	27.07	0	1.414	74
9	8	225	20	0	0	76
10	10	225	20	0	0	79
11	9	225	20	0	0	83
12	2	225	20	0	0	81

Table 2.8 presents the data resulting from an investigation into the effect of two variables, reaction temperature ( $\xi_1$ ), and reactant concentration ( $\xi_2$ ), on the percentage conversion of a chemical process ( $y$ ). The process engineers had used an approach to improving this process based on designed experiments. The first experiment was a screening experiment involving several factors that isolated temperature and concentration as the two most important variables. Because the experimenters thought that the process was operating in the vicinity of the optimum, they elected to fit a quadratic model relating yield to temperature and concentration.

Panel A of Table 2.8 shows the levels used for  $\xi_1$  and  $\xi_2$  in the natural units of measurements. Panel B shows the levels in terms of coded variables  $x_1$  and  $x_2$ . Figure 2.8 shows



**Figure 2.8** Central composite design for the chemical process example.

the experimental design in Table 2.8 graphically. This design is called a **central composite design**, and it is widely used for fitting a second-order response surface. Notice that the design consists of four runs at the corners of a square, plus four runs at the center of this square, plus four **axial runs**. In terms of the coded variables the corners of the square are  $(x_1, x_2) = (-1, -1), (1, -1), (-1, 1), (1, 1)$ ; the center points are at  $(x_1, x_2) = (0, 0)$ ; and the axial runs are at  $(x_1, x_2) = (-1.414, 0), (1.414, 0), (0, -1.414), (0, 1.414)$ .

We will fit the second-order model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \varepsilon$$

using the coded variables. The matrix  $\mathbf{X}$  and vector  $\mathbf{y}$  for this model are

$$\mathbf{X} = \begin{array}{c} \begin{array}{cccccc} & x_1 & x_2 & x_1^2 & x_2^2 & x_1 x_2 \\ \begin{bmatrix} 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1.414 & 0 & 2 & 0 & 0 \\ 1 & 1.414 & 0 & 2 & 0 & 0 \\ 1 & 0 & -1.414 & 0 & 2 & 0 \\ 1 & 0 & 1.414 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} & \mathbf{y} = \begin{bmatrix} 43 \\ 78 \\ 69 \\ 73 \\ 48 \\ 76 \\ 65 \\ 74 \\ 76 \\ 79 \\ 83 \\ 81 \end{bmatrix} \end{array} \end{array}$$

Notice that we have shown the variables associated with each column above that column in the matrix  $\mathbf{X}$ . The entries in the columns associated with  $x_1^2$  and  $x_2^2$  are found by squaring the entries in columns  $x_1$  and  $x_2$ , respectively, and the entries in the  $x_1 x_2$  column are found by multiplying each entry from  $x_1$  by the corresponding entry from  $x_2$ . The matrix  $\mathbf{X}'\mathbf{X}$  and vector  $\mathbf{X}'\mathbf{y}$  are

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 12 & 0 & 0 & 8 & 8 & 0 \\ 0 & 8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 8 & 0 & 0 & 0 \\ 8 & 0 & 0 & 12 & 4 & 0 \\ 8 & 0 & 0 & 4 & 12 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}, \quad \mathbf{X}'\mathbf{y} = \begin{bmatrix} 845.000 \\ 78.592 \\ 33.726 \\ 511.000 \\ 541.000 \\ -31.000 \end{bmatrix}$$

and from  $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$  we obtain

$$\mathbf{b} = \begin{bmatrix} 79.75 \\ 10.18 \\ 4.22 \\ -8.50 \\ -5.25 \\ -7.75 \end{bmatrix}$$

Therefore the fitted model for percentage conversion is

$$\hat{y} = 79.75 + 10.18x_1 + 4.22x_2 - 8.50x_1^2 - 5.25x_2^2 - 7.75x_1x_2$$

In terms of the natural variables, the model is

$$\hat{y} = -1080.22 + 7.7671\xi_1 + 23.1932\xi_2 - 0.0136\xi_1^2 - 0.2100\xi_2^2 - 0.0620\xi_1\xi_2$$

Tables 2.9 and 2.10 present some of the output from Design-Expert, a popular statistics software package for experimental design and RSM. In Table 2.9, the model-fitting procedure is presented. Design-Expert attempts to fit a series of models to the data (first-order or linear, first-order with interaction, second-order, and cubic). The significance of each additional group of terms is assessed sequentially using the partial  $F$ -test discussed previously. Lack-of-fit tests and several summary statistics are reported for each of the models considered. Because the central composite design cannot support a full cubic model, the results for that model are accompanied by a statement “aliased,” indicating that not all of

**TABLE 2.9 Model Fitting for the Chemical Process Example (From Design-Expert)**

<i>Sequential Model Sum of Squares [Type I]</i>						
Source	Sum of Squares	df	Mean Square	$F$ -Value	$P$ -Value Prob > $F$	
Mean vs Total	59784.08	1	59784.08			
Linear vs Mean	970.98	2	485.49	5.30	0.0301	
2FI vs Linear	240.25	1	240.25	3.29	0.1071	
Quadratic vs 2FI	546.42	2	273.21	43.98	0.0003	Suggested
Cubic vs Quadratic	10.02	2	5.01	0.74	0.5346	Aliased
Residual	27.25	4	6.81			
Total	61579.00	12	5131.58			
<i>Lack-of-Fit Tests</i>						
Linear	797.19	6	132.86	14.90	0.0246	
2FI	556.94	5	111.39	12.49	0.0319	
Quadratic	10.52	3	3.51	0.39	0.7682	Suggested
Cubic	0.50	1	0.50	0.056	0.8281	Aliased
Pure Error	26.75	3	8.92			
<i>Model Summary Statistics</i>						
Source	Std. Dev.	$R$ -Squared	Adjusted $R$ -Squared	Predicted $R$ -Squared	PRESS	
Linear	9.57	0.5410	0.4390	0.1814	1469.26	
2FI	8.54	0.6748	0.5529	0.2901	1274.27	
Quadratic	2.49	0.9792	0.9619	0.9318	122.37	Suggested
Cubic	2.61	0.9848	0.9583	0.9557	79.56	Aliased

“*Sequential Model Sum of Squares [Type I]*”: Select the highest order polynomial where the additional terms are significant and the model is not aliased.

“*Lack-of-Fit Tests*”: Want the selected model to have insignificant lack-of-fit.

“*Model Summary Statistics*”: Focus on the model maximizing the “Adjusted  $R$ -Squared” and the “Predicted  $R$ -Squared.”

**TABLE 2.10 Design-Expert ANOVA for the Chemical Process Example**

<i>ANOVA for Response Surface Quadratic Model Analysis of Variance Table [Partial Sum of Squares—Type III]</i>					
Source	Sum of Squares	df	Mean Square	F-Value	P-Value Prob > F
Model	1757.65	5	351.53	56.59	<0.0001
<i>A-Temp</i>	828.78	1	828.78	133.42	<0.0001
<i>B-Conc</i>	142.20	1	142.20	22.89	0.0030
<i>AB</i>	240.25	1	240.25	38.68	0.0008
<i>A</i> <sup>2</sup>	462.40	1	462.40	74.44	0.0001
<i>B</i> <sup>2</sup>	176.40	1	176.40	28.40	0.0018
Residual	37.27	6	6.21		
<i>Lack of Fit</i>	10.52	3	3.51	0.39	0.7682
<i>Pure Error</i>	26.75	3	8.92		
Cor Total	1794.92	11			
Std. Dev.	2.49			R-Squared	0.9792
Mean	70.58			Adj R-Squared	0.9619
C.V. %	3.53			Pred R-Squared	0.9318
PRESS	122.37			Adeq. Precision	20.367

Factor	Coefficient Estimate	Standard df	95% CI		
			Error	Low	High
Intercept	79.75	1	1.25	76.70	82.80
<i>A-Temp</i>	10.18	1	0.88	8.02	12.33
<i>B-Conc</i>	4.22	1	0.88	2.06	6.37
<i>AB</i>	-7.75	1	1.25	-10.80	-4.70
<i>A</i> <sup>2</sup>	-8.50	1	0.99	-10.91	-6.09
<i>B</i> <sup>2</sup>	-5.25	1	0.99	-7.66	-2.84

*Final Equation in Terms of Coded Factors:*

Conversion	=	
+79.75		
+10.18	*A	
+4.22	*B	
-7.75	*A *B	
-8.50	*A <sup>2</sup>	
-5.25	*B <sup>2</sup>	

*Final Equation in Terms of Actual Factors:*

Conversion	=	
-1080.21867		
+7.76713	*Temp	
+23.19320	*Conc	
-0.062000	*Temp *Conc	
-0.013600	*Temp <sup>2</sup>	
-0.21000	*Conc <sup>2</sup>	

the cubic model parameters can be uniquely estimated. The summary statistics strongly support selecting the quadratic model, and this is also the model selected by Design-Expert.

To illustrate the computations in Table 2.9, consider testing the significance of the linear (first-order) terms. The regression sum of squares for this model, that is, the contribution of  $x_1$  and  $x_2$  over the mean (intercept) is  $SS_R(\beta_1, \beta_2|\beta_0) = 970.98$ . The residual mean square for the linear model can be shown to be  $MS_E = 91.58$ . Therefore, the  $F$ -statistic for testing the contribution of the linear terms beyond the intercept is

$$F_0 = \frac{SS_R(\beta_1, \beta_2|\beta_0)/2}{MS_E} = \frac{970.98/2}{91.58} = \frac{485.99}{91.58} = 5.30$$

for which the  $P$ -value is  $P = 0.0301$ , so we conclude that the linear terms are significant. The analysis in Table 2.9 indicates that the linear, interaction and pure quadratic terms contribute significantly to the model. Table 2.10 summarizes the ANOVA for the quadratic model. In this table, the source of variation labeled "Lack of Fit" contains all of the sums of squares for terms higher-order than quadratic. Table 2.10 also reports 95% confidence intervals for each model parameter. None of these CIs include zero, indicating that there are no nonsignificant terms in the model. If some of these CIs had included zero, some analysts would drop the nonsignificant variables for the model, resulting in a reduced quadratic model for the process. It is also possible to employ variable selection methods such as stepwise regression to determine the subset of variables to include in the model. Generally, we prefer to fit the full quadratic model whenever possible, unless there are large differences between the full and the reduced model in PRESS and adjusted  $R^2$ . Table 2.10 indicates that the  $R^2$  and adjusted  $R^2$  values for this model are satisfactory. The  $R^2_{\text{prediction}}$  based on PRESS is

$$\begin{aligned} R^2_{\text{prediction}} &= 1 - \frac{\text{PRESS}}{SS_T} \\ &= 1 - \frac{122.37}{1799.92} \\ &= 0.9318 \end{aligned}$$

indicating that the model will probably explain a high percentage (about 93%) of the variability in new data.

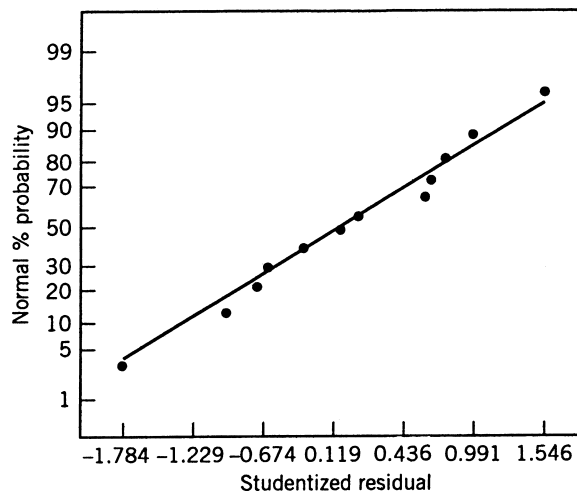
Table 2.11 contains the observed and predicted values of percentage conversion, the residuals, and other diagnostic statistics for this model. None of the studentized residuals or the values of  $R$ -student are large enough to indicate any problem with outliers. Notice that the hat diagonals  $h_{ii}$  take on only two values, either 0.625 or 0.250. The values of  $h_{ii} = 0.625$  are associated with the four runs at the corners of the square in the design and the four axial runs. All eight of these points are equidistant from the center of the design; this is why all of the  $h_{ii}$  values are identical. The four center points all have  $h_{ii} = 0.250$ . Figures 2.9, 2.10, and 2.11 show a normal probability plot of the studentized residuals, a plot of the studentized residuals versus the predicted values  $\hat{y}_i$ , and a plot of the studentized residuals versus run order. None of these plots reveal any model inadequacy.

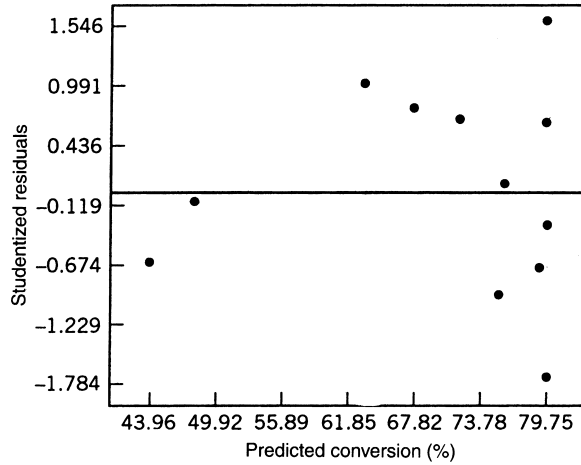
Plots of the conversion response surface and the contour plot, respectively, for the fitted model are shown in Fig. 2.12a and b. The response surface plots indicate that the maximum percentage conversion is at about 240°C and 20% concentration.

**TABLE 2.11** Observed Values, Predicted Values, Residuals, and Other Diagnostics for the Chemical Process Example

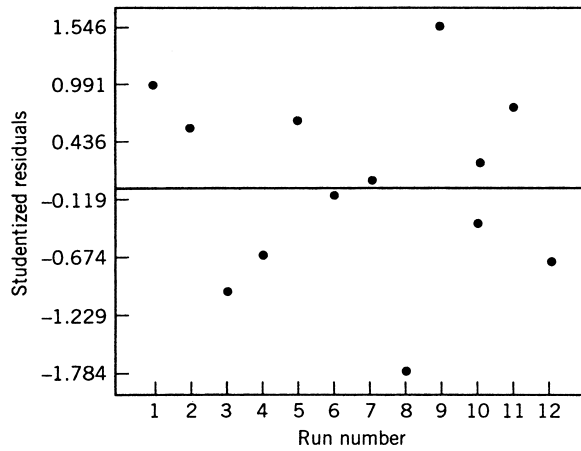
Obs. Order	Actual Value	Predicted Value	Residual	$h_{ii}$	Student Residual	Cook's $D$	$R$ -Student
1	43.00	43.96	-0.96	0.625	-0.643	0.115	-0.609
2	78.00	79.11	-1.11	0.625	-0.745	0.154	-0.714
3	69.00	67.89	1.11	0.625	0.748	0.155	0.717
4	73.00	72.04	0.96	0.625	0.646	0.116	0.612
5	48.00	48.11	-0.11	0.625	-0.073	0.001	-0.067
6	76.00	75.90	0.10	0.625	-0.073	0.001	-0.067
7	65.00	63.54	1.46	0.625	0.982	0.268	0.979
8	74.00	75.46	-1.46	0.625	-0.985	0.269	-0.982
9	76.00	79.75	-3.75	0.250	-1.784	0.177	-2.377
10	79.00	79.75	-0.75	0.250	-0.357	0.007	-0.329
11	83.00	79.75	3.25	0.250	1.546	0.133	1.820
12	81.00	79.75	1.25	0.250	0.595	0.020	0.560

In many response surface problems the experimenter is interested in predicting the response  $y$  or estimating the mean response at a particular point in the process variable space. The response surface plots in Fig. 2.12 give a graphical display of these quantities. Typically, the variance of the prediction is also of interest, because this is a direct measure of the likely uncertainty associated with the point estimate produced by the model. Recall from Equation 2.40 that the variance of the estimate of the mean response at the point  $\mathbf{x}_0$  is given by  $\text{Var}[\hat{y}_i(\mathbf{x}_0)] = \sigma^2 \mathbf{x}'_0(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0$ . Plots of the estimated standard deviation of the predicted response,  $\sqrt{\text{Var}[\hat{y}_i(\mathbf{x}_0)]}$ , obtained by estimating  $\sigma^2$  by the mean square error  $MS_E = 5.89$  for this model for all values of  $\mathbf{x}_0$  in the region of experimentation, are presented in Fig. 2.13a and b. Both the response surface in Fig. 2.13a and the contour plot

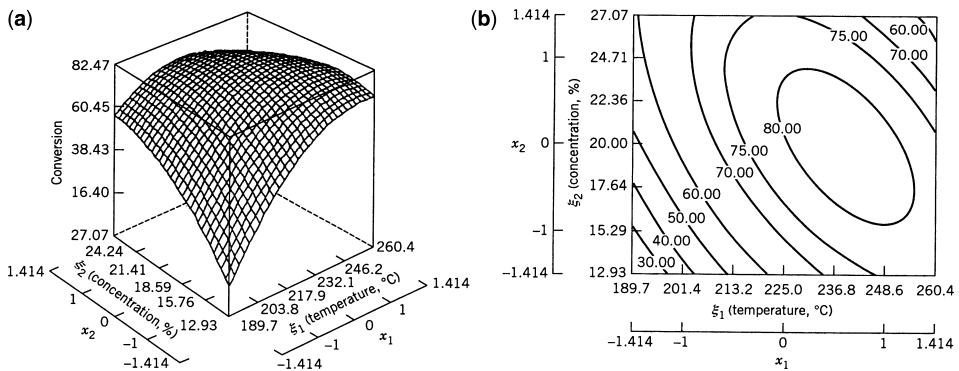
**Figure 2.9** Normal probability plot of the studentized residuals, chemical process example.



**Figure 2.10** Plot of studentized residuals versus predicted conversion, chemical process example.

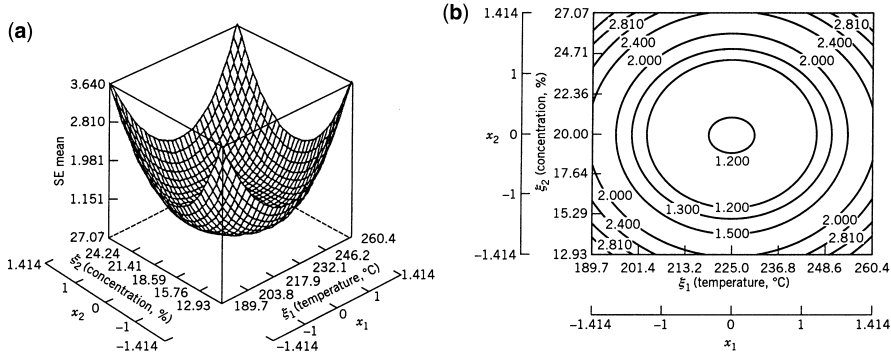


**Figure 2.11** Plot of the studentized residuals versus run order, chemical process example.



**Figure 2.12** (a) Response surface of predicted conversion. (b) Contour plot of predicted conversion.





**Figure 2.13** (a) Response surface plot of  $\sqrt{\widehat{\text{Var}}[\hat{y}(x_0)]}$ . (b) Contour plot of  $\sqrt{\widehat{\text{Var}}[\hat{y}(x_0)]}$ .

of constant  $\sqrt{\widehat{\text{Var}}[\hat{y}(x_0)]}$  in Fig. 2.13b show that the  $\sqrt{\widehat{\text{Var}}[\hat{y}(x_0)]}$  is the same for all points  $\mathbf{x}_0$  that are the same distance from the center of the design. This is a result of the spacing of the axial runs in the central composite design at 1.414 units from the origin (in the coded variables), and is a design property called **rotatability**. This is an important property for a second-order response surface design, and it will be discussed in more detail in Chapter 8.

JMP will also produce excellent results when fitting a second-order model. Table 2.12 is a portion of the JMP output for the chemical process example of this section. The output contains similar information as in the Design-Expert output we saw earlier, but it is organized somewhat differently. JMP will also produce response surface and contour plots similar to those in Figure 2.12, but it also features a **prediction profiler** graph which is the bottom figure in Table 2.12. This graph shows the fitted response surface as seen along each individual coordinate axis (the solid curve), along with the prediction interval on the response (the dotted line curves). This set of curves is interactive. If the crosshairs shown on the plot are dragged to a new location for one variable, say  $x_1$ , the response curve for the other variable will shift due to the interaction between the two factors. For example Figure 2.14 shows the prediction profiler after  $x_1$  is changed to a new value (237.31). Notice the change in the shape of the response surface as a function of  $x_2$ . This is a very useful feature for visualizing the shape of response surfaces in three or more factors.

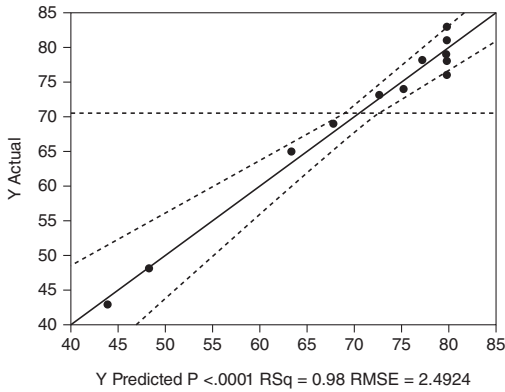
## 2.9 QUALITATIVE REGRESSOR VARIABLES

The regression models employed in RSM usually involved **quantitative** variables—that is, variables that are measured on a numerical scale. For example, variables such as temperature, pressure, distance, and age are quantitative variables. Occasionally, we need to incorporate **qualitative** variables in a regression model. For example, suppose that one of the variables in a regression model is the machine from which each observation  $y_i$  is taken. Assume that only two machines are involved. We may wish to assign different levels to the two machines to allow for the possibility that each machine may have a different effect on the response.

**TABLE 2.12 JMP Output for the Chemical Process Example**

**Response Y**

**Actual by Predicted Plot**



**Summary of Fit**

RSquare	0.979235
RSquare Adj	0.961931
Root Mean Square Error	2.492353
Mean of Response	70.58333
Observations (or Sum Wgts)	12

**Analysis of Variance**

Source	DF	Sum of Squares	Mean Square	F-Ratio
Model	5	1757.6457	351.529	56.5903
Error	6	37.2709	6.212	<b>Prob &gt; F</b>
C. Total	11	1794.9167		<.0001*

**Lack of Fit**

Source	DF	Sum of Squares	Mean Square	F-Ratio
Lack of Fit	3	10.520942	3.50698	0.3933
Pure Error	3	26.750000	8.91667	<b>Prob &gt; F</b>

Source	DF	Sum of Squares	Mean Square	F-Ratio
Total error	6	37.270942		0.7682
				<b>Max RSq</b>

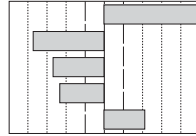
TABLE 2.12 (Continued)

Parameter Estimates

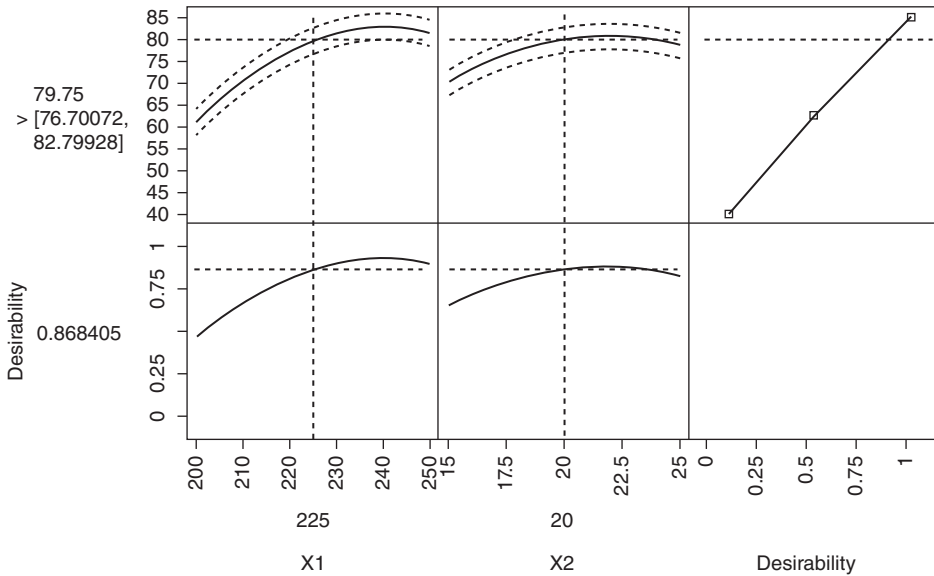
Term	Estimate	Std Error	t-Ratio	Prob > t
Intercept	79.75	1.246177	64.00	<.0001*
X1(200,250)	10.178301	0.88118	11.55	<.0001*
X2(15,25)	4.2159903	0.88118	4.78	0.0030*
X1*X2	-7.75	1.246177	-6.22	0.0008*
X1*X1	-8.5	0.985189	-8.63	0.0001*
X2*X2	-5.25	0.985189	-5.33	0.0018*

Sorted Parameter Estimates

Term	Estimate	Std Error	t-Ratio	Prob > t
X1(200,250)	10.178301	0.88118	11.55	<.000
X1*X1	-8.5	0.985189	-8.63	0.000
X1*X2	-7.75	1.246177	-6.22	0.000
X2*X2	-5.25	0.985189	-5.33	0.001
X2(15,25)	4.2159903	0.88118	4.78	0.003



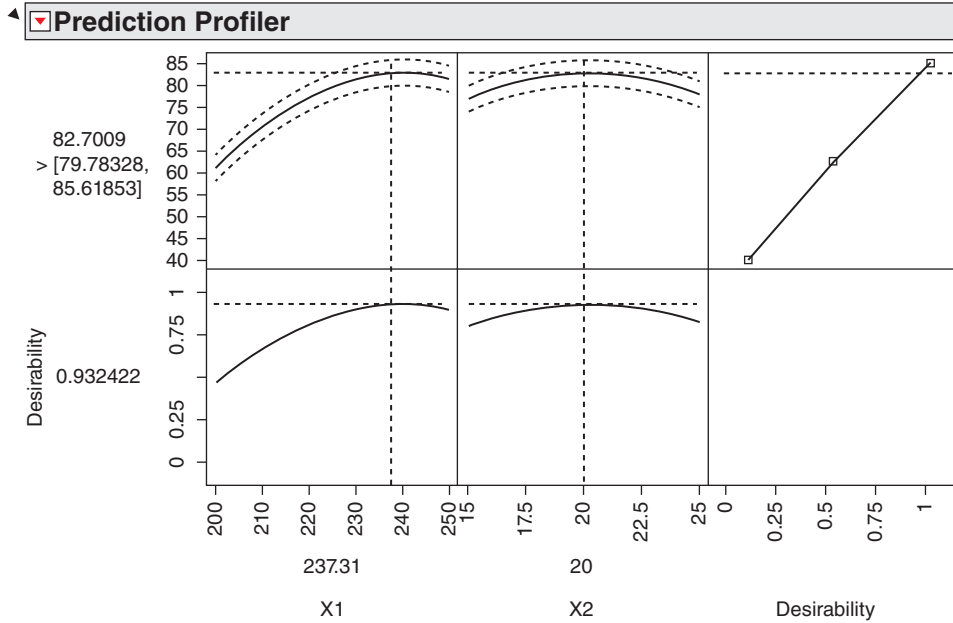
Prediction Profiler



The usual method of representing the different levels of a qualitative variable is by using **indicator variables**. For example, to introduce the effect of two different machines into a regression model, we could define an indicator variable as follows:

$$x = 0 \quad \text{if the observation is from machine 1}$$

$$x = 1 \quad \text{if the observation is from machine 2}$$



**Figure 2.14** The prediction profiler for the chemical process example with  $X_1 = 237.31$ .

In general, a qualitative variable with  $t$  levels is represented by  $t - 1$  indicator variables, which are assigned the values either 0 or 1. Thus, if there were three machines, the different levels would be accounted for by two indicator variables defined as follows:

$x_1$	$x_2$	
0	0	if the observation is from machine 1
1	0	if the observation is from machine 2
0	1	if the observation is from machine 3

Indicator variables are also referred to as **dummy variables**. The following example illustrates some of the uses of indicator variables. For other applications, see Montgomery, Peck, and Vining (2012) and Myers (1990).

**Example 2.11 The Surface Finish Regression Model** A mechanical engineer is investigating the surface finish of metal parts produced on a lathe and its relationship to the speed [in revolutions per minute (RPM)] of the lathe. The data are shown in Table 2.13. Note that the data have been collected using two different types of cutting tools. Because it is likely that the type of cutting tool affects the surface finish, we will fit the model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

**TABLE 2.13 Surface Finish Data for Example 2.11**

Observation Number, $i$	Surface Finish, $y_i$	RPM	Type of Cutting Tool
1	45.44	225	302
2	42.03	200	302
3	50.10	250	302
4	48.75	245	302
5	47.92	235	302
6	47.79	237	302
7	52.26	265	302
8	50.52	259	302
9	45.58	221	302
10	44.78	218	302
11	33.50	224	416
12	31.23	212	416
13	37.52	248	416
14	37.13	260	416
15	34.70	243	416
16	33.92	238	416
17	32.13	224	416
18	35.47	251	416
19	33.49	232	416
20	32.29	216	416

where  $y$  is the surface finish,  $x_1$  is the lathe speed in RPM, and  $x_2$  is an indicator variable denoting the type of cutting tool used; that is,

$$x_2 = \begin{cases} 0 & \text{for tool type 302} \\ 1 & \text{for tool type 416} \end{cases}$$

The parameters in this model may be easily interpreted. If  $x_2 = 0$ , then the model becomes

$$y = \beta_0 + \beta_1 x_1 + \varepsilon$$

which is a straight-line model with slope  $\beta_1$  and intercept  $\beta_0$ . However, if  $x_2 = 1$ , the model becomes

$$y = \beta_0 + \beta_1 x_1 + \beta_2(1) + \varepsilon = (\beta_0 + \beta_2) + \beta_1 x_1 + \varepsilon$$

which is a straight-line model with slope  $\beta_1$  and intercept  $\beta_0 + \beta_2$ . Thus, the model  $y = \beta_0 + \beta_1 x + \beta_2 x_2 + \varepsilon$  implies that surface finish is linearly related to lathe speed and that the slope  $\beta_1$  does not depend on the type of cutting tool used. However, the type of cutting tool does affect the intercept, and  $\beta_2$  indicates the change in the intercept associated with a change in tool type from 302 to 416.

The model matrix  $\mathbf{X}$  and vector  $\mathbf{y}$  for this problem are as follows:

$$\mathbf{X} = \begin{bmatrix} 1 & 225 & 0 \\ 1 & 200 & 0 \\ 1 & 250 & 0 \\ 1 & 245 & 0 \\ 1 & 235 & 0 \\ 1 & 237 & 0 \\ 1 & 265 & 0 \\ 1 & 259 & 0 \\ 1 & 221 & 0 \\ 1 & 218 & 0 \\ 1 & 224 & 0 \\ 1 & 212 & 1 \\ 1 & 248 & 1 \\ 1 & 260 & 1 \\ 1 & 243 & 1 \\ 1 & 238 & 1 \\ 1 & 224 & 1 \\ 1 & 251 & 1 \\ 1 & 232 & 1 \\ 1 & 216 & 1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 45.44 \\ 42.03 \\ 50.10 \\ 48.75 \\ 47.92 \\ 47.79 \\ 52.26 \\ 50.52 \\ 45.58 \\ 44.78 \\ 33.50 \\ 31.23 \\ 37.52 \\ 37.13 \\ 34.70 \\ 33.92 \\ 32.13 \\ 35.47 \\ 33.49 \\ 32.29 \end{bmatrix}$$

The fitted model is

$$\hat{y} = 14.27620 + 0.14115x_1 - 13.28020x_2$$

The analysis of variance for this model is shown in Table 2.14. Note that the hypothesis  $H_0: \beta_1 = \beta_2 = 0$  (significance of regression) is rejected. This table also contains the sums of squares

$$\begin{aligned} SS_R &= SS_R(\beta_1, \beta_2 | \beta_0) \\ &= SS_R(\beta_1 | \beta_0) + SS_R(\beta_2 | \beta_1, \beta_0) \end{aligned}$$

so that a test of the hypothesis  $H_0: \beta_2 = 0$  can be made. This hypothesis is also rejected, so we conclude that tool type has an effect on surface finish.

**TABLE 2.14** Analysis of Variance of Example 2.11

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	$F_0$	$P$ -Value
Regression	1012.0595	1	506.0297	1103.69	$1.0175 \times 10^{-18}$
$SS_R(\beta_1   \beta_0)$	(130.6091)	(1)	130.6091	284.87	$4.6980 \times 10^{-12}$
$SS_R(\beta_2   \beta_1, \beta_0)$	(881.4504)	(1)	881.4504	1922.52	$6.2439 \times 10^{-19}$
Error	7.7943	17	0.4508		
Total	1019.8538	19			

It is also possible to use indicator variables to investigate whether tool type affects both the slope and intercept. Let the model be reformulated as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon$$

where  $x_2$  is the indicator variable. Now if tool type 302 is used, then  $x_2 = 0$ , and the model is

$$y = \beta_0 + \beta_1 x_1 + \varepsilon$$

If tool type 416 is used, then  $x_2 = 1$ , and the model becomes

$$\begin{aligned} y &= \beta_0 + \beta_1 x_1 + \beta_2 + \beta_3 x_1 + \varepsilon \\ &= (\beta_0 + \beta_2) + (\beta_1 + \beta_3) x_1 + \varepsilon \end{aligned}$$

Note that  $\beta_2$  is the change in the intercept, and  $\beta_3$  is the change in slope produced by a change in tool type.

Another method of analyzing this data set is to fit separate regression models to the data for each tool type. However, the indicator variable approach has several advantages. First, only one regression model must be estimated. Second, by pooling the data on both tool types, more degrees of freedom for error are obtained. Third, tests of both hypotheses on the parameters  $\beta_2$  and  $\beta_3$  are just special cases of the extra sum of squares method.

## 2.10 TRANSFORMATION OF THE RESPONSE VARIABLE

We noted in Section 2.7 that often a data transformation can be used when residual analysis indicates some problem with underlying model assumptions, such as nonnormality or nonconstant variance in the response variable. In this section we illustrate the use of data transformation for a classic example from the literature. Following the example, we discuss methods for the choice of the transformation.

**Example 2.12 The Worstest Yarn Data** The data in Table 2.15 [taken from Box and Draper (1987)] show the number of cycles to failure of worsted yarn ( $y$ ) and three factors defined as follows:

$$\begin{aligned} \text{Length of test specimen (mn):} & \quad x_1 = \frac{\text{Length} - 300}{50} \\ \text{Amplitude of load cycle (MM):} & \quad x_2 = \text{Length} - 9 \\ \text{Load (grams):} & \quad x_3 = \frac{\text{Length} - 45}{5} \end{aligned}$$

These factors form a  $3^3$  factorial experiment. This experiment will support a complete second-order polynomial. The least squares fit is

$$\begin{aligned} \hat{y} &= 550.7 + 660x_1 - 535.9x_2 - 310.8x_3 + 238.7x_1^2 + 275.7x_2^2 \\ &\quad - 48.3x_3^2 - 456.5x_1x_2 - 235.7x_1x_3 + 143x_2x_3 \end{aligned}$$

**TABLE 2.15 The Worsted Yarn Data**

Run Number	Length, $x_1$	Amplitude, $x_2$	Load, $x_3$	Cycles to Failure, $y$
1	-1	-1	-1	674
2	0	-1	-1	1414
3	1	-1	-1	3636
4	-1	0	-1	338
5	0	0	-1	1022
6	1	0	-1	1368
7	-1	1	-1	170
8	0	1	-1	442
9	1	1	-1	1140
10	-1	-1	0	370
11	0	-1	0	1198
12	1	-1	0	3184
13	-1	0	0	266
14	0	0	0	620
15	1	0	0	1070
16	-1	1	0	118
17	0	1	0	332
18	1	1	0	884
19	-1	-1	1	292
20	0	-1	1	634
21	1	-1	1	2000
22	-1	0	1	210
23	0	0	1	438
24	1	0	1	566
25	-1	1	1	90
26	0	1	1	220
27	1	1	1	360

The  $R^2$  value is 0.975. An analysis of variance is given in Table 2.16. The fit appears to be reasonable and both the first- and second-order terms appear to be necessary.

Figure 2.15 is a plot of residuals versus the predicted cycles to failure  $\hat{y}_i$  for this model. There is an indication of an outward-opening funnel in this plot, implying possible inequality of variance.

**TABLE 2.16 Analysis of Variance for the Quadratic Model for the Worsted Yarn Data**

Source of Variability	Sum of Squares ( $\times 10^{-3}$ )	Degrees of Freedom	Mean Square ( $\times 10^{-3}$ )	$F_0$
First-order terms	14,748.5	3	4,916.2	70
Added second-order terms	4,224.3	6	704.1	9.5
Residual	1,256.6	17	73.9	
Total	20,229.4	26		



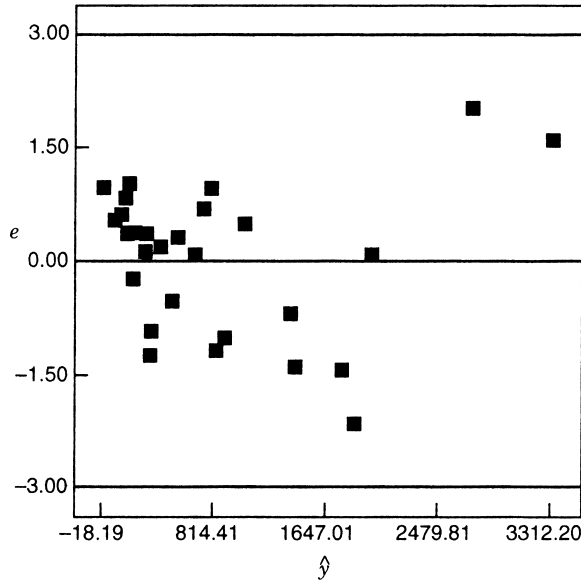


Figure 2.15 Plot of residuals versus predicted for the worsted yarn, experiment, quadratic model.

When a natural log transformation is used for  $y$ , we obtain the following model:

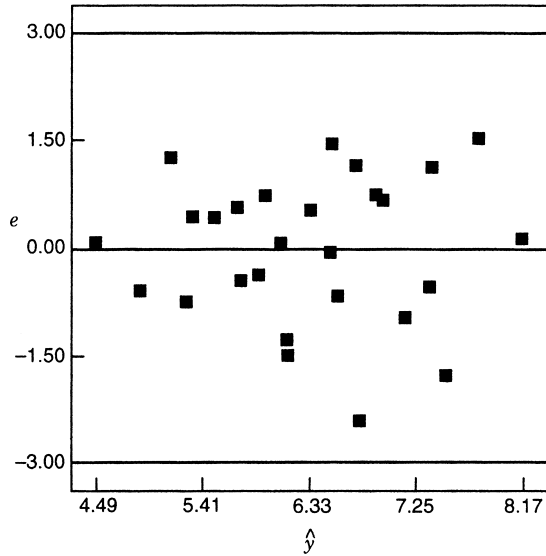
$$\begin{aligned}\ln \hat{y} &= 6.33 + 0.82x_1 - 0.63x_2 - 0.38x_3 \\ \hat{y} &= e^{6.33+0.82x_1-0.63x_2-0.38x_3}\end{aligned}$$

This model has  $R^2 = 0.963$ , and has only three model terms (apart from the intercept). None of the second-order terms are significant. Here, as in most modeling exercises, simplicity is of vital importance. The elimination of the quadratic terms and interaction terms with the change in response metric not only allows a better fit than the second-order model with the natural metric, but the effect of the design variables  $x_1$ ,  $x_2$ , and  $x_3$  on the response is clear.

Figure 2.16 is a plot of residuals versus the predicted response for the log model. There is still some indication of inequality of variance, but the log model, overall, is an improvement on the original quadratic fit.

In the previous example, we illustrated the problem of nonconstant variance in the response variable  $y$  in linear regression and noted that this is a departure from the standard least squares assumptions. This inequality of variance problem occurs fairly often in practice, often in conjunction with a non-normal response variable. Examples include a count of defects or particles, proportion data such as yield or fraction defective, or a response variable that follows some skewed distribution (one tail of the response distribution is longer than the other). We illustrated how **transformation of the response variable** can be used for stabilizing the variance of the response. In our example we selected a log transformation empirically, by noting that it greatly improved the appearance of the residual plots.

Generally, transformations are used for three purposes: stabilizing the response variance, making the distribution of the response variable closer to the normal distribution, and improving the fit of the model to the data. This last objective could include model simplification, say by eliminating interaction, or higher-order polynomial terms. Sometimes



**Figure 2.16** Plot of residuals predicted for the worsted yam data, log model.

a transformation will be reasonably effective in simultaneously accomplishing more than one of these objectives.

We often find that the **power family** of transformations  $y^* = y^\lambda$  is very useful, where  $\lambda$  is the parameter of the transformation to be determined (e.g.,  $\lambda = \frac{1}{2}$  means use the square root of the original response). Box and Cox (1964) have shown how the transformation parameter  $\lambda$  may be estimated simultaneously with the other model parameters (overall mean and treatment effects). The theory underlying their method uses the method of maximum likelihood. The actual computational procedure consists of performing, for various values of  $\lambda$ , a standard analysis of variance on

$$y^{(\lambda)} = \begin{cases} \frac{y^{\lambda-1}}{\lambda \hat{y}^{\lambda-1}} & \lambda \neq 0 \\ \hat{y} \ln y & \lambda = 0 \end{cases} \quad (2.68)$$

where  $\hat{y} = \ln^{-1}[(1/n)\Sigma \ln y]$  is the geometric mean of the observations. The maximum likelihood estimate of  $\lambda$  is the value for which the error sum of squares, say  $SS_E(\lambda)$  is a minimum. This value of  $\lambda$  is usually found by plotting a graph of  $SS_E(\lambda)$  versus  $\lambda$  and then reading the value of  $\lambda$  that minimizes  $SS_E(\lambda)$  from the graph. Usually between 10 and 20 values of  $\lambda$  are sufficient for estimation of the optimum value. A second iteration using a finer mesh of values can be performed if a more accurate estimate of  $\lambda$  is necessary.

Notice that we *cannot* select a value of  $\lambda$  by *directly* comparing the error sums of squares from analyses of variance on  $y^\lambda$ , because for each value of  $\lambda$  the error sum of squares is measured on a different scale. Furthermore, a problem arises in  $y$  when  $\lambda = 0$ ; namely, as  $\lambda$  approaches zero,  $y^\lambda$  approaches unity. That is, when  $\lambda = 0$ , all the response values are a constant. The component  $(y^\lambda - 1)/\lambda$  of Equation 2.68 alleviates this problem because as

$\lambda$  tends to zero,  $(y^\lambda - 1)/\lambda$  goes to a limit of  $\ln y$ . The divisor component  $y^{\lambda-1}$  in Equation 2.68 rescales the responses so that the error sums of squares are directly comparable.

In applying the Box–Cox method, we recommend using simple choices for  $\lambda$ , because the practical difference between  $\lambda = 0.5$  and  $\lambda = 0.58$  is likely to be small, but the square root transformation ( $\lambda = 0.5$ ) is much easier to interpret. Obviously, values of  $\lambda$  close to unity would suggest that no transformation is necessary.

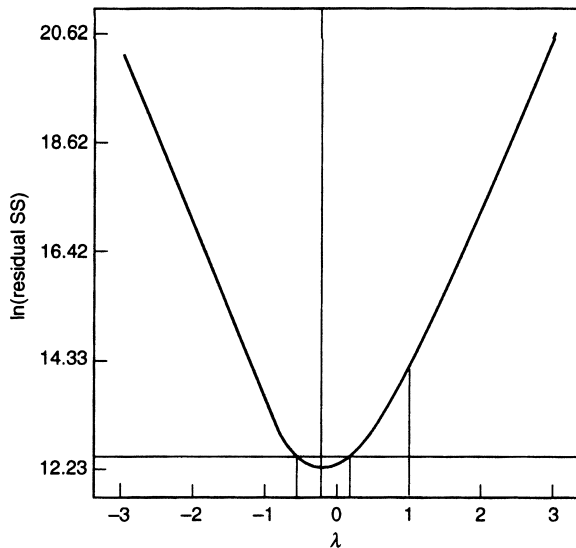
Once a value of  $\lambda$  is selected by the Box–Cox method, the experimenter can analyze the data using  $y^\lambda$  as the response, unless of course  $\lambda = 0$ , in which case use  $\ln y$ . It is perfectly acceptable to use  $y^{(\lambda)}$  as the actual response, although the model parameter estimates will have a scale difference and origin shift in comparison with the results obtained using  $y^\lambda$  (or  $\ln y$ ).

An approximate  $100(1-\alpha)\%$  confidence interval for  $\lambda$  can be found by computing

$$SS^* = SS_E(\lambda) \left( 1 + \frac{t_{\alpha/2, \nu}^2}{\nu} \right) \quad (2.69)$$

where  $\nu$  is the number of degrees of freedom, and plotting a line parallel to the  $\lambda$ -axis at height  $SS^*$  on the graph of  $SS_E(\lambda)$  versus  $\lambda$ . Then by locating the points on the  $\lambda$ -axis where  $SS^*$  cuts the curve  $SS_E(\lambda)$ , we can read confidence limits on  $\lambda$  directly from the graph. If this confidence interval includes the value  $\lambda = 1$ , this implies that the data do not support the need for the transformation.

Several software packages have implemented the Box–Cox procedure. Figure 2.17 shows the output graphics from Design-Expert when the Box–Cox procedure is applied to the worsted yarn data in Table 2.14. The optimum value of  $\lambda$  is  $-0.24$ , and the 95% confidence interval for  $\lambda$  contains zero, so the use of a log transformation is indicated.



**Figure 2.17** The Box–Cox procedure applied to the worsted yarn data in Table 2.5.

## EXERCISES

- 2.1 A study was performed on wear of a bearing  $y$  and its relationship to  $\xi_1$  = oil viscosity and  $\xi_2$  = load. The following data were obtained:

$y$	$\xi_1$	$\xi_2$
193	1.6	851
230	15.5	816
172	22.0	1058
91	43.0	1201
113	33.0	1357
125	40.0	1115

- (a) Fit a multiple linear regression model to these data, using coded variables  $x_1$  and  $x_2$ , defined so that  $-1 \leq x_i \leq 1$ ,  $i = 1, 2$ .
- (b) Convert the model in part (a) to a model in the natural variables  $\xi_1$  and  $\xi_2$ .
- (c) Use the model to predict wear when  $\xi_1 = 25$  and  $\xi_2 = 1000$ .
- (d) Fit a multiple linear regression model with an interaction term to these data. Use the coded variables defined in part (a).
- (e) Use the model in (d) to predict wear when  $\xi_1 = 25$  and  $\xi_2 = 1000$ . Compare this prediction with the predicted value from part (b) above.
- 2.2 Consider the regression models developed in Exercise 2.1.
- (a) Test for significance of regression from the first-order model in part (a) of Exercise 2.1 using the analysis of variance with  $\alpha = 0.05$ . What are your conclusions?
- (b) Use the extra sum of squares method to investigate adding the interaction term to the model [part (d) of Exercise 2.1]. With  $\alpha = 0.05$ , what are your conclusions about this term?
- 2.3 The pull strength of a wire bond is an important characteristic. Table E2.1 gives information on pull strength ( $y$ ), die height ( $x_1$ ), post height ( $x_2$ ), loop height ( $x_3$ ), wire length ( $x_4$ ), bond width on the die ( $x_5$ ), and bond width on the post ( $x_6$ ).
- (a) Fit a multiple linear regression model using  $x_2$ ,  $x_3$ ,  $x_4$ , and  $x_5$  as the regressors.
- (b) Test for significance of regression using the analysis of variance with  $\alpha = 0.05$ . What are your conclusions?
- (c) Use the model from part (a) to predict pull strength when  $x_2 = 20$ ,  $x_3 = 30$ ,  $x_4 = 90$ , and  $x_5 = 2.0$ .
- 2.4 An engineer at a semiconductor company wants to model the relationship between the device gain  $hFE(y)$  and three parameters: emitter RS ( $x_1$ ), base RS ( $x_2$ ), and emitter-to-base RS ( $x_3$ ). The data are shown in Table E2.2.
- (a) Fit a multiple linear regression model to the data.
- (b) Predict  $hFE$  when  $x_1 = 14.5$ ,  $x_2 = 220$ , and  $x_3 = 5.0$ .
- (c) Test for significance of regression using the analysis of variance with  $\alpha = 0.05$ . What conclusions can you draw?

**TABLE E2.1 Wire Bond Data for Exercise 2.3**

$y$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
8.0	5.2	19.6	29.6	94.9	2.1	2.3
8.3	5.2	19.8	32.4	89.7	2.1	1.8
8.5	5.8	19.6	31.0	96.2	2.0	2.0
8.8	6.4	19.4	32.4	95.6	2.2	2.1
9.0	5.8	18.6	28.6	86.5	2.0	1.8
9.3	5.2	18.8	30.6	84.5	2.1	2.1
9.3	5.6	20.4	32.4	88.8	2.2	1.9
9.5	6.0	19.0	32.6	85.7	2.1	1.9
9.8	5.2	20.8	32.2	93.6	2.3	2.1
10.0	5.8	19.9	31.8	86.0	2.1	1.8
10.3	6.4	18.0	32.6	87.1	2.0	1.6
10.5	6.0	20.6	33.4	93.1	2.1	2.1
10.8	6.2	20.2	31.8	83.4	2.2	2.1
11.0	6.2	20.2	32.4	94.5	2.1	1.9
11.3	6.2	19.2	31.4	83.4	1.9	1.8
11.5	5.6	17.0	33.2	85.2	2.1	2.1
11.8	6.0	19.8	35.4	84.1	2.0	1.8
12.3	5.8	18.8	34.0	86.9	2.1	1.8
12.5	5.6	18.6	34.2	83.0	1.9	2.0

**TABLE E2.2 Gain Data for Exercise 2.4**

$x_1$ Emitter RS	$x_2$ Base RS	$x_3$ E-to-B RS	$y$ $hFE$
14.620	226.00	7.000	128.40
15.630	220.00	3.375	52.62
14.620	217.40	6.375	113.90
15.000	220.00	6.000	98.01
14.500	226.50	7.625	139.90
15.250	224.10	6.000	102.60
16.120	220.50	3.375	48.14
15.130	223.50	6.125	109.60
15.500	217.60	5.000	82.68
15.130	228.50	6.625	112.60
15.500	230.20	5.750	97.52
16.120	226.50	3.750	59.06
15.130	226.60	6.125	111.80
15.630	225.60	5.375	89.09
15.380	234.00	8.875	171.90
15.500	230.00	4.000	66.80
14.250	224.30	8.000	157.10
14.500	240.50	10.870	208.40
14.620	223.70	7.375	133.40

**TABLE E2.3 Electric Power Consumption Data for Exercise 2.5**

$y$	$x_1$	$x_2$	$x_3$	$x_4$
240	25	24	91	100
236	31	21	90	95
290	45	24	88	110
274	60	25	87	88
301	65	25	91	94
316	72	26	94	99
300	80	25	87	97
296	84	25	86	96
267	75	24	88	110
276	60	25	91	105
288	50	25	90	100
261	38	23	89	98

- 2.5** The electric power consumed each month by a chemical plant is thought to be related to the average ambient temperature ( $x_1$ ), the number of days in the month ( $x_2$ ), the average product purity ( $x_3$ ), and the tons of product produced ( $x_4$ ). The past year's historical data are available and are presented in Table E2.3.
- (a) Fit a multiple linear regression model to the data.
- (b) Predict power consumption for a month in which  $x_1 = 75^\circ\text{F}$ ,  $x_2 = 24$  days,  $x_3 = 90\%$ , and  $x_4 = 98$  tons.
- 2.6** Heat treating is often used to carburize metal parts, such as gears. The thickness of the carburized layer is considered an important feature of the gear, and it contributes to the overall reliability of the part. Because of the critical nature of this feature, two different lab tests are performed on each furnace load. One test is run on a sample pin that accompanies each load. The other test is a destructive test, where an actual part is cross-sectioned. This test involved running a carbon analysis on the surface of both the gear pitch (top of the gear tooth) and the gear root (between the gear teeth). The data in Table E2.4 are the results of the pitch carbon analysis test for 32 parts.
- (a) Fit a linear regression model relating the results of the pitch carbon analysis test (PITCH) to the five regressor variables.
- (b) Test for significance of regression. Use  $\alpha = 0.05$ .
- 2.7** A regression model  $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \varepsilon$  has been fitted to a sample of  $n = 25$  observations. The calculated  $t$ -ratios  $b_j/se(b_j)$ ,  $j = 1, 2, 3$ , are as follows: for  $\beta_1$ ,  $t_0 = 4.82$ ; for  $\beta_2$ ,  $t_0 = 8.21$ ; and for  $\beta_3$ ,  $t_0 = 0.98$ .
- (a) Find  $P$ -values for each of the  $t$ -statistics.
- (b) Using  $\alpha = 0.05$ , what conclusions can you draw about the regressor  $x_3$ ? Does it seem likely that this regressor contributes significantly to the model?
- 2.8** Consider the electric power consumption data in Exercise 2.5.
- (a) Estimate  $\sigma^2$  for the model fit in Exercise 2.5.
- (b) Use the  $t$ -test to assess the contribution of each regressor to the model. Using  $\alpha = 0.01$ , what conclusions can you draw?

**TABLE E2.4 Heat Treating Data for Exercise 2.6**

TEMP	SOAKTIME	SOAKPCT	DIFFTIME	DIFFPCT	PITCH
1650	0.58	1.10	0.25	0.90	0.013
1650	0.66	1.10	0.33	0.90	0.016
1650	0.66	1.10	0.33	0.90	0.015
1650	0.66	1.10	0.33	0.95	0.016
1600	0.66	1.15	0.33	1.00	0.015
1600	0.66	1.15	0.33	1.00	0.016
1650	1.00	1.10	0.50	0.80	0.014
1650	1.17	1.10	0.58	0.80	0.021
1650	1.17	1.10	0.58	0.80	0.018
1650	1.17	1.10	0.58	0.80	0.019
1650	1.17	1.10	0.58	0.90	0.021
1650	1.17	1.10	0.58	0.90	0.019
1650	1.17	1.15	0.58	0.90	0.021
1650	1.20	1.15	1.10	0.80	0.025
1650	2.00	1.15	1.00	0.80	0.025
1650	2.00	1.10	1.10	0.80	0.026
1650	2.20	1.10	1.10	0.80	0.024
1650	2.20	1.10	1.10	0.80	0.025
1650	2.20	1.15	1.10	0.80	0.024
1650	2.20	1.10	1.10	0.90	0.025
1650	2.20	1.10	1.10	0.90	0.027
1650	2.20	1.10	1.50	0.90	0.026
1650	3.00	1.15	1.50	0.80	0.029
1650	3.00	1.10	1.50	0.70	0.030
1650	3.00	1.10	1.50	0.75	0.028
1650	3.00	1.15	1.66	0.85	0.032
1650	3.33	1.10	1.50	0.80	0.033
1700	4.00	1.10	1.50	0.70	0.039
1650	4.00	1.10	1.50	0.70	0.040
1650	4.00	1.15	1.50	0.85	0.035
1700	12.50	1.00	1.50	0.70	0.056
1700	18.50	1.00	1.50	0.70	0.068

**2.9** Consider the bearing wear data in Exercise 2.1.

- Estimate  $\sigma^2$  for the no-interaction model.
- Compute the  $t$ -statistic for each regression coefficient. Using  $\alpha = 0.05$ , what conclusions can you draw?
- Use the extra sum of squares method to investigate the usefulness of adding  $x_2 =$  load to the model that already contains  $x_1 =$  oil viscosity. Use  $\alpha = 0.05$ .

**2.10** Consider the wire bond pull strength data in Exercise 2.3.

- Estimate  $\sigma^2$  for this model.
- Find the standard errors for each of the regression coefficients.
- Calculate the  $t$ -test statistic for each regression coefficient. Using  $\alpha = 0.05$ , what conclusions can you draw? Do all variables contribute to the model?

- 2.11** Reconsider the semiconductor data in Exercise 2.4.
- Estimate  $\sigma^2$  for the model you have fitted to the data.
  - Find the standard errors of the regression coefficients.
  - Calculate the  $t$ -test statistic for each regression coefficient. Using  $\alpha = 0.05$ , what conclusions can you draw?
- 2.12** Exercise 2.6 presents data on heat treating gears.
- Estimate  $\sigma^2$  for the model.
  - Find the standard errors of the regression coefficients.
  - Evaluate the contribution of each regressor to the model using the  $t$ -test with  $\alpha = 0.05$ .
  - Fit a new model to the response PITCH using new regressors  $x_1 = \text{SOAKTIME} \times \text{SOAKPCT}$  and  $x_2 = \text{DIFFTIME} \times \text{DIFFPCT}$ .
  - Test the model in part (d) for significance of regression using  $\alpha = 0.05$ . Also calculate the  $t$ -test for each regressor and draw conclusions.
  - Estimate  $\sigma^2$  for the model from part (d), and compare this with the estimate of  $\sigma^2$  obtained in part (b) above. Which estimate is smaller? Does this offer any insight regarding which model might be preferable?
- 2.13** Consider the wire bond pull strength data in Exercise 2.3.
- Find 95% confidence intervals on the regression coefficients.
  - Find a 95% confidence interval on mean pull strength when  $x_2 = 20$ ,  $x_3 = 30$ ,  $x_4 = 90$ , and  $x_5 = 2.0$ .
- 2.14** Consider the semiconductor data in Exercise 2.4.
- Find 99% confidence intervals on the regression coefficients.
  - Find a 99% prediction interval on  $hFE$  when  $x_1 = 14.5$ ,  $x_2 = 220$ , and  $x_3 = 5.0$ .
  - Find a 99% confidence interval on mean  $hFE$  when  $x_1 = 14.5$ ,  $x_2 = 220$ , and  $x_3 = 5.0$ .
- 2.15** Consider the heat-treating data from Exercise 2.6.
- Find 95% confidence intervals on the regression coefficients.
  - Find a 95% interval on mean PITCH on  $\text{TEMP} = 1650$ ,  $\text{SOAKTIME} = 1.00$ ,  $\text{SOAKPCT} = 1.10$ ,  $\text{DIFFTIME} = 1.00$ , and  $\text{DIFFPCT} = 0.80$ .
- 2.16** Reconsider the heat treating in Exercise 2.6 and 2.12, where we fit a model to PITCH using regressors  $x_1 = \text{SOAKTIME} \times \text{SOAKPCT}$  and  $x_2 = \text{DIFFTIME} \times \text{DIFFPCT}$ .
- Using the model with regressors  $x_1$  and  $x_2$ , find a 95% confidence interval on mean PITCH when  $\text{SOAKTIME} = 1.00$ ,  $\text{SOAKPCT} = 1.10$ ,  $\text{DIFFTIME} = 1.00$ , and  $\text{DIFFPCT} = 0.80$ .
  - Compare the length of this confidence interval with the length of the confidence interval on mean PITCH at the same point from Exercise 2.15 part (b), where an additive model in SOAKTIME, SOAKPCT, DIFFTIME, and DIFFPCT was used. Which confidence interval is shorter? Does this tell you anything about which model is preferable?



- 2.17** For the regression model for the wire bond pull strength data in Exercise 2.3.
- Plot the residuals versus  $\hat{y}$  and versus the regressors used in the model. What information is provided by these plots?
  - Construct a normal probability plot of the residuals. Are there reasons to doubt the normality assumption for this model?
  - Are there any indications of influential observations in the data?
- 2.18** Consider the semiconductor *hFE* data in Exercise 2.4.
- Plot the residuals from this model versus  $\hat{y}$ . Comment on the information in this plot.
  - What is the value of  $R^2$  for this model?
  - Refit the model using  $\ln hFE$  as the response variable.
  - Plot the residuals versus predicted  $\ln hFE$  for the model in part (c) above. Does this give any information about which model is preferable?
  - Plot the residuals from the model in part (d), versus the regressor  $x_3$ . Comment on this plot.
  - Refit the model to  $\ln hFE$  using  $x_1$ ,  $x_2$ , and  $1/x_3$  as the regressors. Comment on the effect of this change in the model.
- 2.19** Consider the regression model for the heat-treating data in Exercise 2.6.
- Calculate the percentage of variability explained by this model.
  - Construct a normal probability plot for the residuals. Comment on the normality assumption.
  - Plot the residuals versus  $\hat{y}$ , and interpret the display.
  - Calculate Cook's distance for each observation, and provide an interpretation of this statistic.
- 2.20** In Exercise 2.12 we fitted a model to the response PITCH in the heat treating data of Exercise 2.6 using new regressors  $x_1 = \text{SOAKTIME} \times \text{SOAKPCT}$  and  $x_2 = \text{DIFFTIME} \times \text{DIFFPCT}$ .
- Calculate the  $R^2$  for this model, and compare it with the value of  $R^2$  from the original model in Exercise 2.6. Does this provide some information about which model is preferable?
  - Plot the residuals from this model versus  $\hat{y}$  and on a normal probability scale. Comment on model adequacy.
  - Find the values of Cook's distance measure. Are any observations unusually influential?
- 2.21** An article entitled "A Method for Improving the Accuracy of Polynomial Regression Analysis" in the *Journal of Quality Technology* (1971, pp. 149–155) reported the following data on  $y =$  ultimate shear strength of a rubber compound (psi) and  $x =$  cure temperature ( $^{\circ}\text{F}$ ).

$y$	770	800	840	810	735	640	590	560
$x$	280	284	292	295	298	305	308	315

- Fit a second-order polynomial to these data.

- (b) Test for significance of regression, using  $\alpha = 0.05$ .
- (c) Test the hypothesis that  $\beta_{11} = 0$ , using  $\alpha = 0.05$ .
- (d) Compute the residuals and use them to evaluate model adequacy.
- (e) Suppose that the following additional observations are available:  $(x = 284, y = 815)$ ,  $(x = 295, y = 830)$ ,  $(x = 305, y = 660)$ , and  $(x = 315, y = 545)$ . Test the second-order model for lack of fit.

**2.22** Consider the following data, which result from an experiment to determine the effect of  $x =$  test time in hours at a particular temperature to  $y =$  change in oil viscosity:

$y$	-4.42	-1.39	-1.55	-1.89	-2.43	-3.15	-4.05	-5.15	-6.43	-7.89
$x$	0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50

- (a) Fit a second-order polynomial to the data.
  - (b) Test for significance of regression, using  $\alpha = 0.05$ .
  - (c) Test the hypothesis that  $\beta_{11} = 0$ , using  $\alpha = 0.05$ .
  - (d) Compute the residuals and use them to evaluate model adequacy.
- 2.23** When fitting polynomial regression models we often subtract  $\bar{x}$  from each  $x$ -value to produce a **centered** regressor  $x' = x - \bar{x}$ . This reduces the effects of dependences among the model terms and often leads to more accurate estimates of the regression coefficients. Using the data from Exercise 2.21, fit the model  $y = \beta_0^* + \beta_1^* x' + \beta_{11}^* (x')^2 + \varepsilon$ . Use the results to estimate the coefficients in the uncentered model  $y = \beta_0 + \beta_1 x + \beta_{11} x^2 + \varepsilon$ .
- 2.24** Suppose that we use a standardized variable  $x' = (x - \bar{x})/s_x$ , where  $s_x$  is the standard deviation of  $x$ , in constructing a polynomial regression model. Using the data in Exercise 2.21 and the standardized variable approach, fit the model  $y = \beta_0^* + \beta_1^* x' + \beta_{11}^* (x')^2 + \varepsilon$ .
- (a) What value of  $y$  do you predict when  $x = 285^\circ\text{F}$ ?
  - (b) Estimate the regression coefficients in the unstandardized model  $y = \beta_0 + \beta_1 x + \beta_{11} x^2 + \varepsilon$ .
  - (c) What can you say about the relationship between  $SS_E$  and  $R^2$  for the standardized and unstandardized models?
  - (d) Suppose that  $y' = (y - \bar{y})/s_y$  is used in the model along with  $x'$ . Fit the model and comment on the relationship between  $SS_E$  and  $R^2$  in the standardized model and the unstandardized model.
- 2.25** An article in the *Journal of Pharmaceuticals Sciences* (vol. 80, 1991, pp. 971–977) presents data on the observed mole fraction solubility of a solute at a constant temperature to the dispersion, dipolar, and hydrogen bonding Hansen partial solubility parameters. The data are in Table E2.5, where  $y$  is the negative logarithm of the mole fraction solubility,  $x_1$  is the dispersion Hansen partial solubility,  $x_2$  is the dipolar partial solubility, and  $x_3$  is the hydrogen bonding partial solubility.
- (a) Fit the model  $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \varepsilon$ .
  - (b) Test for significance of regression, using  $\alpha = 0.05$ .

TABLE E2.5 Data for Exercise 2.25

Obs	y	$x_1$	$x_2$	$x_3$
1	0.22200	7.3	0.0	0.0
2	0.39500	8.7	0.0	0.3
3	0.42200	8.8	0.7	1.0
4	0.43700	8.1	4.0	0.2
5	0.42800	9.0	0.5	1.0
6	0.46700	8.7	1.5	2.8
7	0.44400	9.3	2.1	1.0
8	0.37800	7.6	5.1	3.4
9	0.49400	10.0	0.0	0.3
10	0.45600	8.4	3.7	4.1
11	0.45200	9.3	3.6	2.0
12	0.11200	7.7	2.8	7.1
13	0.43200	9.8	4.2	2.0
14	0.10100	7.3	2.5	6.8
15	0.23200	8.5	2.0	6.6
16	0.30600	9.5	2.5	5.0
17	0.09230	7.4	2.8	7.8
18	0.11600	7.8	2.8	7.7
19	0.07640	7.7	3.0	8.0
20	0.43900	10.3	1.7	4.2
21	0.09440	7.8	3.3	8.5
22	0.11700	7.1	3.9	6.6
23	0.07260	7.7	4.3	9.5
24	0.04120	7.4	6.0	10.9
25	0.25100	7.3	2.0	5.2
26	0.00002	7.6	7.8	20.7

- (c) Plot the residuals, and comment on model adequacy.  
 (d) Use the extra sum of squares method to test the contribution of the second-order terms, using  $\alpha = 0.05$ .

**2.26** Below are data on  $y =$  green liquor concentration (g/liter) and  $x =$  paper machine speed (ft/min) from a kraft paper machine (the data were read from a graph in an article in the *Tappi Journal*, March, 1986).

y	16.0	15.8	15.6	15.5	14.8	14.0	13.5	13.0	12.0	11.0
x	1700	1720	1730	1740	1750	1760	1770	1780	1790	1795

- (a) Fit the model  $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$  using least squares.  
 (b) Test for significance of regression using  $\alpha = 0.05$ . What are your conclusions?  
 (c) Test the contribution of the quadratic term to the model, over the contribution of the linear term, using an  $F$ -statistic. If  $\alpha = 0.05$ , what conclusion can you draw?  
 (d) Plot the residuals from this model versus  $\hat{y}$ . Does the plot reveal any inadequacies?  
 (e) Construct a normal probability plot of the residuals. Comment on the normality assumption.

- 2.27 Consider a multiple regression model with  $k$  regressors. Show that the test statistic for significance of regression can be written as

$$F_k = \frac{R^2/k}{(1 - R^2)/(n - k - 1)}$$

Suppose that  $n = 20$ ,  $k = 4$ , and  $R^2 = 0.90$ . If  $\alpha = 0.05$ , what conclusion would you draw about the relationship between  $y$  and the four regressors?

- 2.28 A regression model is used to relate a response  $y$  to  $k = 4$  regressors. What is the smallest value of  $R^2$  that will result in a significant regression if  $\alpha = 0.05$ ? Use the results of the previous exercise. Are you surprised by how small the value of  $R^2$  is?
- 2.29 Show that we can express the residuals from a multiple regression model as

$$\mathbf{e} = (\mathbf{I} - \mathbf{H})\mathbf{y}$$

$$\text{where } \mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'.$$

- 2.30 Show that the variance of the  $i$ th residual  $e_i$  in a multiple regression model is  $\sigma^2(1 - h_{ii})$  and that the covariance between  $e_i$  and  $e_j$  is  $-\sigma^2 h_{ij}$ , where the  $h$ s are the elements of  $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ .
- 2.31 Consider the multiple linear regression model  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ . If  $\mathbf{b}$  denotes the least squares estimator of  $\boldsymbol{\beta}$ , show that  $\mathbf{b} = \boldsymbol{\beta} + \mathbf{R}\boldsymbol{\varepsilon}$  where  $\mathbf{R} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ .
- 2.32 **Constrained least squares:** Suppose we wish to find the least squares estimator of  $\boldsymbol{\beta}$  in the model  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$  subject to a set of equality constraints, say  $\mathbf{T}\boldsymbol{\beta} = \mathbf{c}$ . Show that the estimator is

$$\mathbf{b}_c = \mathbf{b}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{T}'[\mathbf{T}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{T}']^{-1}(\mathbf{c} - \mathbf{T}\mathbf{b})$$

where  $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ .

- 2.33 In Section 2.10 we introduced transformations as a technique to handle certain types of model inadequacy, including inequality of variance. Another technique, weighted least squares, can also deal with inequality of variance. Suppose that the variance of the  $i$ th observation is  $\sigma_i^2$  but that the errors in the model are uncorrelated. Weighted least squares estimates the model parameters using the objective function

$$L = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{W}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

where  $\mathbf{W}$  is a diagonal matrix of “weights”  $w_i = 1/\sigma_i^2$ .

- (a) Find the weight least squares estimator of the vector of model parameters  $\boldsymbol{\beta}$ .
- (b) Find the covariance matrix of the weighted least squares estimator from part (a).
- 2.34 A hospital conducts surveys of patients when they are discharged to determine their satisfaction with the care and service they have received. Table E2.6 contains a recent sample of this data. The Surgical–Medical variable is an indicator variable, taking on the value 0 if the patient is on the surgical service and 1 if the patient is on the medical service. The Severity and Anxiety variables are determined based on

**TABLE E2.6 Patient Satisfaction Data**

Observation	Age	Severity	Surgical–Medical	Anxiety	Satisfaction
1	55	50	0	2.1	68
2	46	24	1	2.8	77
3	30	46	1	3.3	96
4	35	48	1	4.5	80
5	59	58	0	2.0	43
6	61	60	0	5.1	44
7	74	65	1	5.5	26
8	38	42	1	3.2	88
9	27	42	0	3.1	75
10	51	50	1	2.4	57
11	53	38	1	2.2	56
12	41	30	0	2.1	88
13	37	31	0	1.9	88
14	24	34	0	3.1	102
15	42	30	0	3.0	88
16	50	48	1	4.2	70
17	58	61	1	4.6	52
18	60	71	1	5.3	43
19	62	62	0	7.2	46
20	68	38	0	7.8	56
21	70	41	1	7.0	59
22	79	66	1	6.2	26
23	63	31	1	4.1	52
24	39	42	0	3.5	83
25	49	40	1	2.1	75

examination of the patient upon admission, and increase as either anxiety or severity of illness increase.

- (a) Fit a linear regression model to this data.
  - (b) Based on the test statistics for the fitted model are all regressor variables necessary? If not, fit an appropriate model to the data.
  - (c) Plot the model residuals and comment on model adequacy.
- 2.35** The quality of Pinot Noir wine is thought to be related to the properties of clarity, aroma, body, flavor, and oakiness. Data for 38 wines are given in Table E2.7.
- (a) Fit a multiple linear regression model relating wine quality to these predictors. Do not include the “Region” variable in the model.
  - (b) Test for significance of regression. What conclusions can you draw?
  - (c) Plot the residuals and comment on model adequacy.
  - (d) Discuss the contribution of each regressor to the model. Would a model with fewer regressors be satisfactory?
- 2.36** Reconsider the wine quality data in Table E2.7. The “Region” predictor refers to three distinct geographical regions where the wine was produced. Note that this is a categorical variable.

**TABLE E2.7 Wine Quality Data (found in Minitab)**

Clarity, $x_1$	Aroma, $x_2$	Body, $x_3$	Flavor, $x_4$	Oakiness, $x_5$	Quality, $y$	Region
1	3.3	2.8	3.1	4.1	9.8	1
1	4.4	4.9	3.5	3.9	12.6	1
1	3.9	5.3	4.8	4.7	11.9	1
1	3.9	2.6	3.1	3.6	11.1	1
1	5.6	5.1	5.5	5.1	13.3	1
1	4.6	4.7	5	4.1	12.8	1
1	4.8	4.8	4.8	3.3	12.8	1
1	5.3	4.5	4.3	5.2	12	1
1	4.3	4.3	3.9	2.9	13.6	3
1	4.3	3.9	4.7	3.9	13.9	1
1	5.1	4.3	4.5	3.6	14.4	3
0.5	3.3	5.4	4.3	3.6	12.3	2
0.8	5.9	5.7	7	4.1	16.1	3
0.7	7.7	6.6	6.7	3.7	16.1	3
1	7.1	4.4	5.8	4.1	15.5	3
0.9	5.5	5.6	5.6	4.4	15.5	3
1	6.3	5.4	4.8	4.6	13.8	3
1	5	5.5	5.5	4.1	13.8	3
1	4.6	4.1	4.3	3.1	11.3	1
0.9	3.4	5	3.4	3.4	7.9	2
0.9	6.4	5.4	6.6	4.8	15.1	3
1	5.5	5.3	5.3	3.8	13.5	3
0.7	4.7	4.1	5	3.7	10.8	2
0.7	4.1	4	4.1	4	9.5	2
1	6	5.4	5.7	4.7	12.7	3
1	4.3	4.6	4.7	4.9	11.6	2
1	3.9	4	5.1	5.1	11.7	1
1	5.1	4.9	5	5.1	11.9	2
1	3.9	4.4	5	4.4	10.8	2
1	4.5	3.7	2.9	3.9	8.5	2
1	5.2	4.3	5	6	10.7	2
0.8	4.2	3.8	3	4.7	9.1	1
1	3.3	3.5	4.3	4.5	12.1	1
1	6.8	5	6	5.2	14.9	3
0.8	5	5.7	5.5	4.8	13.5	1
0.8	3.5	4.7	4.2	3.3	12.2	1
0.8	4.3	5.5	3.5	5.8	10.3	1
0.8	5.2	4.8	5.7	3.5	13.2	1

The wine here is Pinot Noir. Region refers to distinct geographic regions.

- (a) Fit the model using the “Region” variable as it is given in Table E2.7. What potential difficulties could be introduced by including this variable in the regression model using the three levels shown in Table E2.7?
- (b) An alternative way to include the categorical variable “Region” would be to introduce two indicator variables  $x_1$  and  $x_2$  as follows:

Region	$x_1$	$x_2$
1	0	0
2	1	0
3	0	1

Why is this approach better than just using the codes 1, 2, and 3?

- (c) Rework Exercise 2.35 using the indicator variables defined in part b for the variable “Region.”

**2.37** Consider the simple linear regression model

$$y = \beta_0 + \beta_1 x + \varepsilon$$

where the variance of the errors  $\varepsilon_i$  is proportional to  $x_i^2$ ; that is,  $\text{Var}(\varepsilon_i^2) = \sigma^2 x_i^2$ .

- (a) Suppose we use the transformations  $y' = y/x$  and  $x' = 1/x$ . Does this stabilize the variance?
- (b) What are the relationships between the parameters in the original and transformed models?
- (c) Suppose that we use weighted least squares technique introduced in Exercise 2.33 with weights  $w_i = 1/x_i^2$ . Is this equivalent to the transformation suggested in part (a)?
- 2.38** Suppose that we want to fit a linear regression model without an intercept term, say  $y_i = \beta x_i + \varepsilon$ . Find a general formula for the least squares estimator of the slope.

### Short Answer Questions

**2.39** Consider the computer output from a regression model shown below.

Source	Sum of Squares	Degrees of Freedom	Mean Square	$F$
Model	532, 567.14	2		
Error				
Total	539, 534.33	14		

- (a) Fill in the blanks in the table.
- (b) How many regressor or predictor variables are in this model?
- (c) Find bounds on the P-value for the test on significance of regression.
- (d) Suppose that the  $t$ -statistic for one of the predictor variables in this model is 4.22. Find bounds on the  $P$ -value.

2.40 Consider the computer output from a regression model shown below.

Source	Sum of Squares	Degrees of Freedom	Mean Square	$F$
Model	534.66		133.67	
Error				
Total	590.25	19		

- (a) Fill in the blanks in the table.
- (b) How many regressor or predictor variables are in this model?
- (c) Compute the  $F$ -statistic for significance of regression and find bounds on the  $P$ -value.
- (d) Suppose that the  $t$ -statistic for one of the predictor variables in this model is 6.11. Find bounds on the  $P$ -value.
- (e) What is the value of the  $R^2$  statistic for this model?
- 2.41 A regression model with 5 regressors has been built to 75 observations. The individual  $t$ -statistics for each predictor are as follows: 4.21, 2.12, 6.98, 3.55, and 2.45.
- (a) Find bounds on the  $P$ -values for each predictor.
- (b) Based on your answers to part (a), are all of the regressor variables currently in the model necessary?
- 2.42 A regression model with 3 regressors has been fit to a sample of 45 observations. The total sum of squares for the model is 275.60 and the model sum of squares is 245.86.
- (a) Find the value of the mean square error.
- (b) Find the model mean square.
- (c) Compute the value of the  $F$ -statistic for significance of regression and find bounds on the  $P$ -value. What are your conclusions?
- (d) What is the value of  $R^2$  for this model?
- (e) What is the value of the adjusted  $R^2$  statistic for this model?
- 2.43 A regression model with 4 regressors has been fit to a sample of 65 observations. The total sum of squares for the model is 300 and the model sum of squares is 280.
- (a) What is the value of  $R^2$  for this model?
- (b) What is the value of the adjusted  $R^2$  statistic for this model?
- (c) Suppose that one of the regressors is removed from the model and the new model sum of squares is 250. What is the value of the adjusted  $R^2$  for this reduced model? What is the impact on model fit of removing this regressor?
- 2.44 A regression model with 3 regressors has been built to a sample of 30 observations. The individual  $t$ -statistics for each predictor are as follows: 2.12, 5.45, and 17.50.
- (a) Find bounds on the  $P$ -values for each predictor.
- (b) Based on your answers to part (a), are all of the regressor variables currently in the model necessary?



- 2.45 The value of the adjusted  $R^2$  statistic always increases when a new regressor variable is added to the model. **True False**
- 2.46 The value of the ordinary  $R^2$  statistic can decrease when a new regressor variable is added to the model. **True False**
- 2.47 The ordinary  $R^2$  statistic is a good indicator of the prediction capability of a regression model. **True False**
- 2.48 If the  $P$ -value for the test for significance of regression is  $< 0.01$ , this is an indication that all of the regressor variables in the model are necessary. **True False**
- 2.49 If the  $P$ -value for the test for significance of regression is  $< 0.01$ , this is an indication that the regression model will be a good predictor of new observations. **True False**
- 2.50 A small value of the PRESS statistic is a good indication that the regression model will be a good predictor of new observations. **True False**



---

# 3

---

## TWO-LEVEL FACTORIAL DESIGNS

### 3.1 INTRODUCTION

Factorial designs are widely used in experiments involving several factors where it is necessary to investigate the joint effects of the factors on a response variable. By joint factor effects, we typically mean main effects and interactions. A very important special case of the factorial design is that where each of the  $k$  factors of interest has only two levels. Because each replicate of such a design has exactly  $2^k$  experimental trials or runs, these designs are usually called  **$2^k$  factorial designs**. These designs are the subject of this chapter.

The  $2^k$  factorial designs are very important in response surface work. Specifically, they find applications in three areas:

1. A  $2^k$  design (or a fractional version discussed in the next chapter) is useful at the start of a response surface study where **screening experiments** should be performed to identify the important process or system variables.
2. A  $2^k$  design is often used to fit a first-order response surface model and to generate the factor effect estimates required to perform the method of **steepest ascent**.
3. The  $2^k$  design is a basic building block used to create other response surface designs. For example, if you augment a  $2^2$  design with axial runs and center points as in Fig. 2.8 of Chapter 2, then a **central composite design** results. Historically, the central composite design is one of the most important designs for fitting second-order response surface models.

### 3.2 THE $2^2$ DESIGN

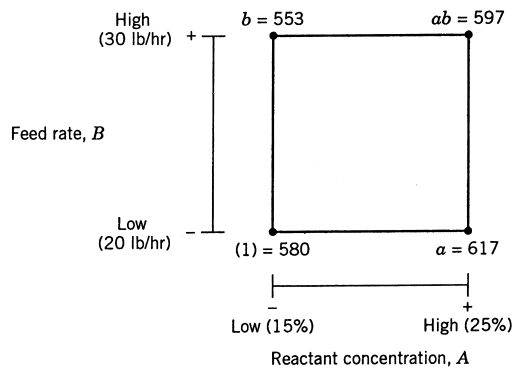
The simplest design in the  $2^k$  series is one with only two factors, say  $A$  and  $B$ , each run at two levels. This design is called a  **$2^2$  factorial design**. The levels of the factors may be arbitrarily called “low” and “high.” These two levels may be **quantitative**, such as two values of temperature or pressure; or they may be **qualitative**, such as two machines, or two operators. In most response surface studies the factors and their levels are quantitative.

As an example, consider an investigation into the effect of the concentration of the reactant and the feed rate on the viscosity of product from a chemical process. Let the reactant concentration be factor  $A$ , and let the two levels of interest be 15% and 25%. The feed rate is factor  $B$ , with the low level being 20 lb/hr and the high level being 30 lb/hr. The experiment is replicated four times, and the data are as follows:

Factor-Level Combination	Viscosity				Total
	I	II	III	IV	
$A$ low, $B$ low	145	148	147	140	580
$A$ high, $B$ low	158	152	155	152	617
$A$ low, $B$ high	135	138	141	139	553
$A$ high, $B$ high	150	152	146	149	597

The four factor-level or treatment combinations in this design are shown graphically in Fig. 3.1. By convention, we denote the effect of a factor by the same capital Latin letter. That is,  $A$  refers to the effect of factor  $A$ ,  $B$  refers to the effect of factor  $B$ , and  $AB$  refers to the  $AB$  interaction. In the  $2^2$  design the low and high levels of  $A$  and  $B$  are denoted by  $-$  (or  $-1$ ) and  $+$  (or  $+1$ ), respectively, on the  $A$ - and  $B$ -axes. Thus,  $-$  on the  $A$ -axis represents the low level of concentration (15%), and  $+$  represents the high level (25%). Similarly,  $-$  on the  $B$ -axis represents the low level of feed rate, and  $+$  denotes the high level.

The four factor-level combinations in the design can also be represented by lowercase letters, as shown in Fig. 3.1. We see from the figure that the high level of any factor at a point in the design is denoted by the corresponding lowercase letter and that the low level of a factor at a point in the design is denoted by the absence of the corresponding letter. Thus,  $a$  represents the combination of factor levels with  $A$  at the high level and  $B$  at the low level,  $b$  represents  $A$  at the low level and  $B$  at the high level, and  $ab$  represents both factors



**Figure 3.1** The  $2^2$  design.

at the high level. By convention, (1) is used to denote the run where both factors are at the low level. This notation is used throughout the 2<sup>k</sup> series.

The **average effect of a factor** (or the **main effect**) is defined as the change in response produced by a change in the level on that factor averaged over the levels of the other factor. Also the symbols (1), *a*, *b*, and *ab* now represent the **totals** of all *n* replicates taken at the points in the design, as illustrated in Fig. 3.1. The formulas for the effects of *A*, *B*, and *AB* may be easily derived. The main effect of *A* can be found as the difference in the average response of the two points on the right-hand side of the square in Fig. 3.1 (call this average  $\bar{y}_{A+}$ , because it is the average response at the points where *A* is at the high level) and the two points on the left-hand side (or  $\bar{y}_{A-}$ ). That is,

$$\begin{aligned} A &= \bar{y}_{A+} - \bar{y}_{A-} \\ &= \frac{ab + a}{2n} - \frac{b + (1)}{2n} \\ &= \frac{1}{2n}[ab + a - b - (1)] \end{aligned} \quad (3.1)$$

The main effect of factor *B* is found as the difference between the average of the response at the two points on the top of the square ( $\bar{y}_{B+}$ ) and the average of the response at the two points on the bottom ( $\bar{y}_{B-}$ ), or

$$\begin{aligned} B &= \bar{y}_{B+} - \bar{y}_{B-} \\ &= \frac{ab + b}{2n} - \frac{a + (1)}{2n} \\ &= \frac{1}{2n}[ab + b - a - (1)] \end{aligned} \quad (3.2)$$

Finally, the interaction effect *AB* is the average of the responses on the right-to-left diagonal points in the square [*ab* and (1)] minus the average of the responses on the left-to-right diagonal points (*a* and *b*), or

$$\begin{aligned} AB &= \frac{ab + (1)}{2n} - \frac{a + b}{2n} \\ &= \frac{1}{2n}[ab + (1) - a - b] \end{aligned} \quad (3.3)$$

Using the example data in Fig. 3.1, we may estimate the effects as

$$\begin{aligned} A &= \frac{1}{2(4)}(597 + 617 - 553 - 580) = 10.125 \\ B &= \frac{1}{2(4)}(597 + 553 - 617 - 580) = -5.875 \\ AB &= \frac{1}{2(4)}(597 + 580 - 617 - 553) = 0.875 \end{aligned}$$

The main effect of *A* (reactant concentration) is positive; this suggests that increasing *A* from the low level (15%) to the high level (25%) will increase the viscosity. The main effect of *B* (feed rate) is negative; this suggests that increasing the feed rate will decrease the viscosity. The interaction effect appears to be small relative to the two main effects.

In many experiments involving  $2^k$  designs, we will examine the magnitude and direction of the factor effects to determine which variables are likely to be important. The **analysis of variance** can generally be used to confirm this interpretation.

In the  $2^k$  design, there are some special, efficient methods for performing the calculations in the analysis of variance. Consider the sums of squares for the factors  $A$ ,  $B$ , and  $AB$ . Note from Equation 3.1 that a **contrast** is used in estimating  $A$ , namely,

$$\text{Contrast}_A = ab + a - b - (1) \quad (3.4)$$

We usually call this contrast the **total effect** of  $A$ . From Equations 3.2 and 3.3, we see that contrasts are also used to estimate  $B$  and  $AB$ . Furthermore, these three contrasts are **orthogonal**. The sum of squares for any contrast is equal to the contrast squared divided by the number of observations in each total in the contrast times the sum of the squares of the contrast coefficients. Consequently, we have

$$SS_A = \frac{[ab + a - b - (1)]^2}{n \times 4} \quad (3.5)$$

$$SS_B = \frac{[ab + b - a - (1)]^2}{n \times 4} \quad (3.6)$$

and

$$\frac{SS_{AB}}{n \times 4} = [ab + (1) - a - b]^2 \quad (3.7)$$

as the sums of squares for  $A$ ,  $B$ , and  $AB$ .

Using the data in Fig. 3.1, we may find the sums of the squares from Equations 3.5, 3.6, and 3.7 as

$$SS_A = \frac{(81)^2}{4(4)} = 410.0625$$

$$SS_B = \frac{(47)^2}{4(4)} = 138.0625$$

and

$$SS_{AB} = \frac{(7)^2}{4(4)} = 3.0625$$

The total sum of squares is found in the usual way; that is,

$$SS_T = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^n y_{ijk}^2 - \frac{y_{\dots}^2}{4n} \quad (3.8)$$

In general,  $SS_T$  has  $4n - 1$  degrees of freedom. The error sum of squares, with  $4(n - 1)$

**TABLE 3.1 Analysis of Variance for Data in Fig. 3.1**

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	$F_0$	$P$ -Value
<i>A</i>	410.0625	1	410.0625	49.08	$1.42 \times 10^{-5}$
<i>B</i>	138.0625	1	138.0625	16.53	0.0016
<i>AB</i>	3.0625	1	3.0625	0.37	0.5562
Error	100.2500	12	8.3542		
Total	651.4375	15			

degrees of freedom, is usually computed by subtraction as

$$SS_E = SS_T - SS_A - SS_B - SS_{AB} \quad (3.9)$$

For the data in Fig. 3.1, we obtain

$$\begin{aligned} SS_T &= \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^4 y_{ijk}^2 - \frac{y_{\dots}^2}{4(4)} \\ &= 334,527.0000 - 344,275.5625 = 651.4375 \end{aligned}$$

and

$$\begin{aligned} SS_E &= SS_T - SS_A - SS_B - SS_{AB} \\ &= 651.4375 - 410.0625 - 138.0625 - 3.0625 \\ &= 100.2500 \end{aligned}$$

using  $SS_A$ ,  $SS_B$ , and  $SS_{AB}$  computed previously. The complete analysis of variance is summarized in Table 3.1. Both main effects are statistically significant, and the interaction is not significant. This confirms our initial interpretation of the data based on the magnitudes of the factor effects.

It is often convenient to write down the factor level or treatment combinations in the order (1), *a*, *b*, *ab*, as in the column labeled “treatment combination” in Table 3.2. This is referred to as **standard order**. The columns labeled *A* and *B* in this table contain the + and – signs corresponding to the four runs at the corners of the square in Fig. 3.1. Note that the contrast coefficients for estimating the interaction effect are just the products of the corresponding coefficients for the two main effects. The column labeled *I* in Table 3.2 represents the total or average of all of the observations in the entire experiment. The column corresponding to *I* has only plus signs and is sometimes called the identity column. The row designators are

**TABLE 3.2 Signs for Calculating Effects in the 2<sup>2</sup> Design**

Treatment Combination	Factorial Effect			
	<i>I</i>	<i>A</i>	<i>B</i>	<i>AB</i>
(1)	+	–	–	+
<i>a</i>	+	+	–	–
<i>b</i>	+	–	+	–
<i>ab</i>	+	+	+	+

the treatment combinations. To find the contrast for estimating any effect, simply multiply the signs in the appropriate column of the table by the corresponding treatment combination and add. For example, to estimate  $A$ , the contrast is  $-(1) + a - b + ab$ , which agrees with Equation 3.1.

**The Regression Model** It is easy to convert the effect estimates in a  $2^k$  factorial design into a regression model that can be used to predict the response at any point in the space spanned by the factors in the design. For the chemical process experiment, the first-order regression model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \quad (3.10)$$

where  $x_1$  is the **coded variable** that represents the reactant concentration,  $x_2$  is the coded variable that represents the feed rate, and the  $\beta$ s are the regression coefficients. The relationship between the natural variables, the reactant concentration and the feed rate, and the coded variables is

$$x_1 = \frac{\text{Conc} - (\text{Conc}_{\text{low}} + \text{Conc}_{\text{high}})/2}{(\text{Conc}_{\text{high}} - \text{Conc}_{\text{low}})/2}$$

and

$$x_2 = \frac{\text{Feed} - (\text{Feed}_{\text{low}} + \text{Feed}_{\text{high}})/2}{(\text{Feed}_{\text{high}} - \text{Feed}_{\text{low}})/2}$$

When the natural variables have only two levels, this coding will produce the  $\pm 1$  notation for the levels of the coded variables. To illustrate this for our example, note that

$$\begin{aligned} x_1 &= \frac{\text{Conc} - (15 + 25)/2}{(25 - 15)/2} \\ &= \frac{\text{Conc} - 20}{5} \end{aligned}$$

Thus, if the concentration is at the high level ( $\text{Conc} = 25\%$ ), then  $x_1 = +1$ , whereas if the concentration is at the low level ( $\text{Conc} = 15\%$ ), then  $x_1 = -1$ . Furthermore,

$$\begin{aligned} x_2 &= \frac{\text{Feed} - (20 + 30)/2}{(30 - 20)/2} \\ &= \frac{\text{Feed} - 25}{5} \end{aligned}$$

Thus, if the feed rate is at the high level ( $\text{Feed} = 30 \text{ lb/hr}$ ), then  $x_2 = +1$ , whereas if the feed rate is at the low level ( $\text{Feed} = 20 \text{ lb/hr}$ ), then  $x_2 = -1$ .

The fitted regression model is

$$\begin{aligned} \hat{y} &= 146.6875 + \left(\frac{10.125}{2}\right)x_1 + \left(\frac{-5.875}{2}\right)x_2 \\ &= 146.6875 + 5.0625x_1 - 2.9375x_2 \end{aligned}$$



where the intercept is the grand average of all 16 observations, and the regression coefficients  $b_1$  and  $b_2$  are one-half the corresponding factor effect estimates. The reason that the regression coefficient is one-half the effect estimate is that a regression coefficient measures the effect of a unit change in  $x$  on the mean of  $y$ , and the effect estimate is based on a two-unit change (from  $-1$  to  $+1$ ). This model explains about 84% of the variability in viscosity, because

$$\begin{aligned} R^2 &= SS_{\text{model}}/SS_T = (SS_A + SS_B)/SS_T \\ &= (410.0625 + 138.0625)/651.4375 \\ &= 548.1250/651.4375 = 0.8414 \end{aligned}$$

This regression model is really a first-order response surface model for the process. If the interaction term in the analysis of variance (Table 3.1) were significant, the regression model would have been

$$y = \beta_0 + \beta_1 x_2 + \beta_2 x_1 + \beta_{12} x_1 x_2 + \varepsilon$$

and the estimate of the regression coefficient  $\beta_{12}$  would be one-half the  $AB$  interaction effect, or  $b_{12} = 0.875/2$ . Clearly this coefficient is small relative to  $b_1$  and  $b_2$ , and so this is further evidence that interaction contributes very little to explaining viscosity in this process.

**Regression Coefficients are Least Squares Estimates** The regression coefficient estimates obtained above are **least squares estimates**. To verify this, write out the model in Equation 3.10 in matrix form, that is,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where

$$\mathbf{y} = \begin{bmatrix} 145 \\ 148 \\ 147 \\ 140 \\ 158 \\ 152 \\ 155 \\ 152 \\ 135 \\ 138 \\ 141 \\ 139 \\ 150 \\ 152 \\ 146 \\ 149 \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \\ \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{14} \\ \varepsilon_{15} \\ \varepsilon_{16} \end{bmatrix}$$

The least squares estimates of  $\beta_0$ ,  $\beta_1$ , and  $\beta_2$  are found from

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

or

$$\begin{aligned} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} &= \begin{bmatrix} 16 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 16 \end{bmatrix}^{-1} \begin{bmatrix} 2347 \\ 81 \\ -47 \end{bmatrix} \\ &= \frac{1}{16} \mathbf{I}_3 \begin{bmatrix} 2347 \\ 81 \\ -47 \end{bmatrix} \end{aligned}$$

where  $\mathbf{I}_3$  is a  $3 \times 3$  identity matrix. This last expression reduces to

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 2347/16 \\ 81/16 \\ -47/16 \end{bmatrix}$$

or  $b_0 = 2347/16 = 146.6875$ ,  $b_1 = 5.0625$ , and  $b_2 = -2.9375$ . Therefore, the least squares estimates of  $\beta_1$  and  $\beta_2$  are exactly one-half of the factor effect estimates. This will always be the case in any  $2^k$  factorial design with coded factors between  $-1$  and  $1$ .

**Residual Analysis** It is easy to compute the **residuals** from a  $2^k$  design via the regression model. For the chemical process experiment, the regression model is

$$\hat{y} = 146.6875 + \left(\frac{10.125}{2}\right)x_1 - \left(\frac{5.875}{2}\right)x_2$$

This model can be used to generate the predicted values of  $y$  at the four points in the design. For example, when the reactant concentration is at the low level ( $x_1 = -1$ ) and the feed rate is at the low level ( $x_2 = -1$ ) the predicted viscosity is

$$\begin{aligned} \hat{y} &= 146.6875 + \left(\frac{10.125}{2}\right)(-1) - \left(\frac{5.875}{2}\right)(-1) \\ &= 144.563 \end{aligned}$$

There are four observations at this treatment combination, and the residuals are

$$\begin{aligned} e_1 &= 145 - 144.563 = 0.438 \\ e_2 &= 148 - 144.563 = 3.438 \\ e_3 &= 152 - 144.563 = 2.438 \\ e_4 &= 140 - 144.563 = -4.563 \end{aligned}$$

The remaining predicted values and residuals are calculated similarly. For the high level of the reactant concentration and the low level of the feed rate:

$$\begin{aligned}\hat{y} &= 146.6875 + \left(\frac{10.125}{2}\right)(+1) + \left(\frac{-5.875}{2}\right)(-1) \\ &= 154.688\end{aligned}$$

and

$$e_5 = 158 - 154.688 = -3.688$$

$$e_6 = 152 - 154.688 = -2.688$$

$$e_7 = 155 - 154.688 = -0.313$$

$$e_8 = 152 - 154.688 = -2.688$$

For the low level of the reactant concentration and the high level of the feed rate:

$$\begin{aligned}\hat{y} &= 146.6875 + \left(\frac{10.125}{2}\right)(-1) + \left(\frac{-5.875}{2}\right)(+1) \\ &= 138.688\end{aligned}$$

and

$$e_9 = 135 - 138.688 = -3.688$$

$$e_{10} = 138 - 138.688 = -0.688$$

$$e_{11} = 141 - 138.688 = 2.313$$

$$e_{12} = 139 - 138.688 = 0.313$$

Finally, for the high level of both factors:

$$\begin{aligned}\hat{y} &= 146.6875 + \left(\frac{10.125}{2}\right)(+1) + \left(\frac{-5.875}{2}\right)(+1) \\ &= 148.812\end{aligned}$$

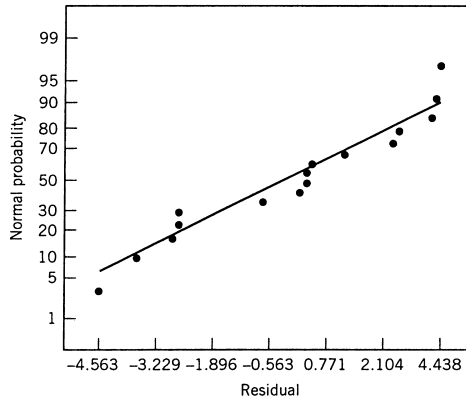
and

$$e_{13} = 150 - 148.812 = 1.188$$

$$e_{14} = 152 - 148.812 = 3.188$$

$$e_{15} = 146 - 148.812 = -2.813$$

$$e_{16} = 149 - 148.812 = 0.188$$

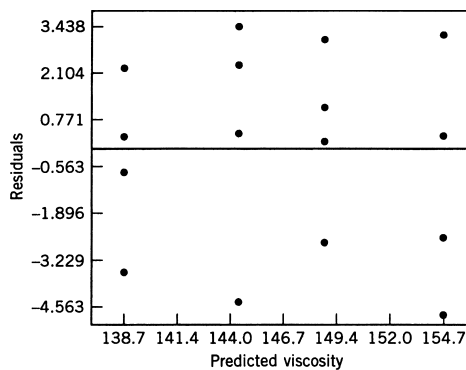


**Figure 3.2** Normal probability plot of the residuals from the chemical process experiment.

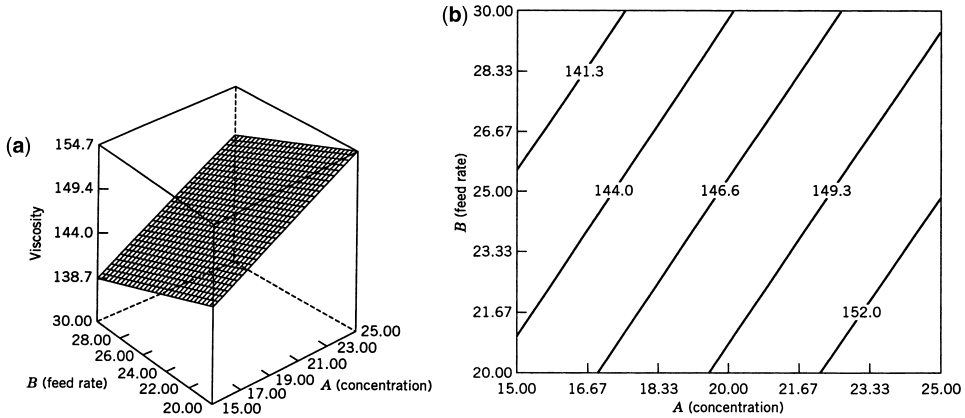
Figure 3.2 presents a normal probability plot of these residuals, and Fig. 3.3 is a plot of the residuals versus the predicted values  $\hat{y}$ . These plots appear satisfactory, so we have no reason to suspect problems with the validity of our conclusions.

**The Response Surface** Figure 3.4a presents a three-dimensional response surface plot of viscosity, based on the first-order model using reactant concentration and feed rate as the regressors, and Fig. 3.4b is the contour plot. Because the model is first-order, the fitted response surface is a plane. Based on examination of this fitted surface, it is clear that viscosity increases as reactant concentration increases and feed rate decreases. Often a fitted surface such as this can be used to determine an appropriate **direction of potential improvement** for a process. A formal method for doing this, called the **method of steepest ascent**, will be introduced in Chapter 5.

**Computer Output** There are several software packages that will construct experimental designs and analyze the resulting data. In this book, we primarily use JMP and Design-Expert. The output from Design-Expert is shown in Table 3.3. In the upper portion of the table an analysis of variance for the full model is presented. The format of this presentation



**Figure 3.3** Plot of residual versus predicted viscosity for the chemical process experiment.



**Figure 3.4** (a) Response surface plot of viscosity from the chemical process experiment. (b) The contour plot.

is slightly different from the analysis of variance in Table 3.1. The source called the “model” is an analysis of variance for the full model using

$$\begin{aligned} SS_{\text{model}} &= SS_A + SS_B + SS_{AB} \\ &= 410.06 + 138.06 + 3.06 \\ &= 551.19 \end{aligned}$$

Thus the statistic

$$F_0 = \frac{MS_{\text{model}}}{MS_E} = \frac{183.73}{8.35} = 21.99$$

is testing the hypotheses

$$\begin{aligned} H_0: & \beta_1 = \beta_2 = \beta_{12} = 0 \\ H_1: & \text{at least one } \beta \neq 0. \end{aligned}$$

Because  $F_0$  is large, we conclude that at least one variable has a nonzero effect. Then each individual factorial effect is tested for significance using the  $F$ -statistic (as in Table 3.1).

The lower portion of Table 3.3 shows the analysis of variance summary for the **reduced** model; that is, the model with the nonsignificant  $AB$  interactions removed. The **error** or **residual** sum of squares is now composed of a **pure error** component arising from the replicate runs and a **lack-of-fit** component corresponding to the  $AB$  interaction. The regression model in terms of both actual and coded variables is given, along with confidence intervals on each model coefficient, a summary of several regression diagnostics from Chapter 2, and the model residuals. The column labeled “Outlier  $t$ ” is  $R$ -student.

**TABLE 3.3 Design Expert Output for the 2<sup>2</sup> Factorial Example**

Response: Viscosity						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F-Value	Prob > F	
Model	551.19	3	183.73	21.99	<0.0001	
A	410.06	1	410.06	49.08	<0.0001	
B	138.06	1	138.06	16.53	0.0016	
AB	3.06	1	3.06	0.37	0.5562	
Pure Error	100.25	12	8.35			
Cor Total	651.44	15				
Std. Dev.	2.89		R-Squared	0.8461		
Mean	146.69		Adj R-Squared	0.8076		
C.V.	1.97		Pred R-Squared	0.7264		
PRESS	178.22		Adeq Precision	11.071		
Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	146.69	1	0.72	145.11	148.26	
A-Reactant C	5.06	1	0.72	3.49	6.64	1.00
B-Feed Rate	-2.94	1	0.72	-4.51	-1.36	1.00
AB	0.44	1	0.72	-1.14	2.01	1.00
Reduced Model:						
Response: Viscosity						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F-Value	Prob > F	
Model	548.13	2	274.06	34.49	<0.0001	
A	410.06	1	410.06	51.60	<0.0001	
B	138.06	1	138.06	17.37	0.0011	
Residual	103.31	13	7.95			
Lack of Fit	3.06	1	3.06	0.37	0.5562	
Pure Error	100.25	12	8.35			
Cor Total	651.44	15				
Std. Dev.	2.82		R-Squared	0.8414		
Mean	146.69		Adj R-Squared	0.8170		

C.V. 1.92 Pred R-Squared 0.7598  
 PRESS 156.50 Adeq Precision 13.107

Factor	Coefficients Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	146.69	1	0.70	145.16	148.21	
A-Reactant C	5.06	1	0.70	3.54	6.59	1.00
B-Feed Rate	-2.94	1	0.70	-4.46	-1.41	1.00

Final Equation in Terms of Coded Factors:

$$\text{Viscosity} = +146.69 + 5.06 * A - 2.94 * B$$

Final Equation in Terms of Actual Factors:

$$\text{Viscosity} = +141.12500 + 1.01250 * \text{Reactant C} - 0.58750 * \text{Feed Rate}$$

Diagnostics Case Statistics

Run Order	Standard Order	Actual Value	Predicted Value	Residual	Leverage	Student Residual	Cook's Distance	Outlier <i>t</i>
1	1	145.00	144.56	0.44	0.188	0.172	0.002	0.166
16	2	148.00	144.56	3.44	0.188	1.353	0.141	1.402
4	3	147.00	144.56	2.44	0.188	0.959	0.071	0.956
5	4	140.00	144.56	-4.56	0.188	-1.796	0.248	-1.989
3	5	158.00	154.69	3.31	0.188	1.304	0.131	1.343
11	6	152.00	154.69	-2.69	0.188	-1.058	0.086	-1.063
8	7	155.00	154.69	0.31	0.188	0.123	0.001	0.118
12	8	152.00	154.69	-2.69	0.188	-1.058	0.086	-1.063
9	9	135.00	138.69	-3.69	0.188	-1.451	0.162	-1.523
14	10	138.00	138.69	-0.69	0.188	-0.271	0.006	-0.261
7	11	141.00	138.69	2.31	0.188	0.910	0.064	0.904
2	12	139.00	138.69	0.31	0.188	0.123	0.001	0.118
15	13	150.00	148.81	1.19	0.188	0.467	0.017	0.453
13	14	152.00	148.81	3.19	0.188	1.254	0.121	1.285
6	15	146.00	148.81	-2.81	0.188	-1.107	0.094	-1.117
10	16	149.00	148.81	0.19	0.188	0.074	0.000	0.071

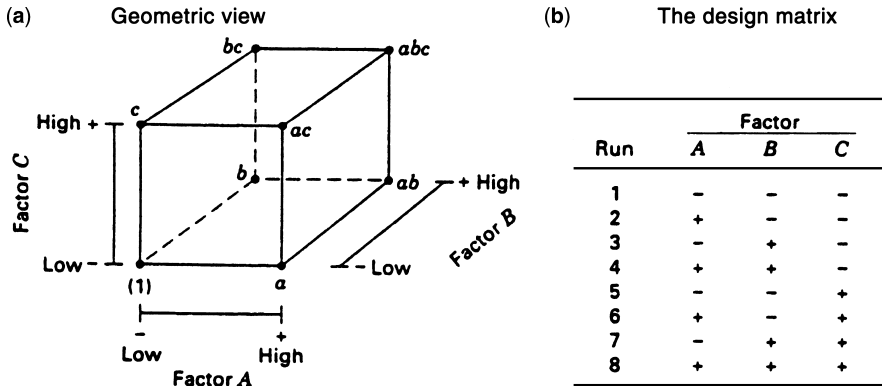


Figure 3.5 The 2<sup>3</sup> factorial design.

### 3.3 THE 2<sup>3</sup> DESIGN

Suppose that three factors, *A*, *B*, and *C*, each at two levels, are of interest. Then the design is called a 2<sup>3</sup> factorial design, and the eight treatment combinations can be displayed graphically as a cube, as shown in Fig. 3.5a. Extending the notation discussed in Section 3.2, we write the treatment combinations in standard order as (1), *a*, *b*, *ab*, *c*, *ac*, *bc*, and *abc*. Remember that these symbols also represent the total of all *n* observations taken at that particular treatment combinations.

Using the “+ and -” notation to represent the high and low levels of each factor, we may list the eight runs in the 2<sup>3</sup> design in the tabular format shown in Fig. 3.5b. This is usually called the **design matrix**.

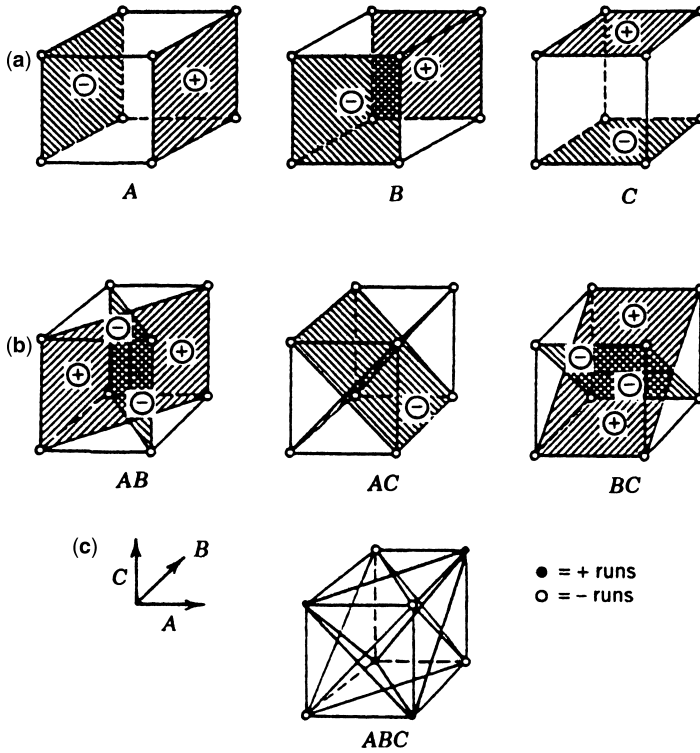
The “+ and -” notation is often called the **geometric notation**. While other notational schemes could be used, we prefer the geometric notation because it facilitates the translation of the analysis of variance results into a regression model. This notation is widely used in response surface methodology.

There are seven degrees of freedom between the eight treatment combinations in the 2<sup>3</sup> design. Three degrees of freedom are associated with the main effects of *A*, *B*, and *C*. Four degrees of freedom are associated with interactions: one each with *AB*, *AC*, and *BC*, and one with *ABC*.

Consider estimating the main effects. First, consider estimating the main effect of *A*. This effect estimate can be expressed as a contrast between the four treatment combinations in the right face of the cube in Fig. 3.6a (where *A* is at the high level) and the four in the left face (where *A* is at the low level). That is, the *A* effect is just the average of the four runs where *A* is at the high level ( $\bar{y}_{A+}$ ) minus the average of the four runs where *A* is at the low level ( $\bar{y}_{A-}$ ), or

$$\begin{aligned}
 A &= \bar{y}_{A+} - \bar{y}_{A-} \\
 &= \frac{a + ab + ac + abc}{4n} - \frac{(1) + b + c + bc}{4n}
 \end{aligned}$$





**Figure 3.6** Geometric view of contrasts corresponding to the main effects and interaction in the 2<sup>3</sup> design. (a) Main effects. (b) Two-factor interactions. (c) Three-factor interaction.

This equation can be rearranged as

$$A = \frac{1}{4n}[a + ab + ac + abc - (1) - b - c - bc] \quad (3.11)$$

In a similar manner, the effect of  $B$  is the difference in averages between the four treatment combinations in the front face of the cube and the four in the back. This yields

$$\begin{aligned} B &= \bar{y}_{B+} - \bar{y}_{B-} \\ &= \frac{1}{4n}[b + ab + bc + abc - (1) - a - c - ac] \end{aligned} \quad (3.12)$$

The effect of  $C$  is the difference in averages between the four treatment combinations in the top face of the cube and the four in the bottom; that is,

$$\begin{aligned} C &= \bar{y}_{C+} - \bar{y}_{C-} \\ &= \frac{1}{4n}[c + ac + bc + abc - (1) - a - b - ab] \end{aligned} \quad (3.13)$$

The two-factor interaction effects may be computed easily. A measure of the  $AB$  interaction is the difference between the average  $A$  effects at the two levels of  $B$ . By convention, one-half of this difference is called the  $AB$  interaction. This is expressed symbolically as follows:

$B$	Average $A$ Effect
High (+)	$\frac{(abc - bc) + (ab - b)}{2n}$
Low (-)	$\frac{(ac - c) + [a - (1)]}{2n}$
Difference	$\frac{abc - bc + ab - b - ac + c - a + (1)}{2n}$

Because the  $AB$  interaction is one-half of this difference, we have

$$AB = \frac{[abc - bc + ab - b - ac + c - a + (1)]}{4n} \tag{3.14}$$

We can write Equation 3.14 as follows:

$$AB = \frac{abc + ab + c + (1)}{4n} - \frac{bc + b + ac + a}{4n}$$

In this form, the  $AB$  interaction is easily seen to be the difference in averages between runs on two diagonal planes in the cube in Fig. 3.6b. Using similar logic and referring to Fig. 3.6b, the  $AC$  and  $BC$  interactions are

$$AC = \frac{1}{4n} [(1) - a + b - ab - c + ac - bc + abc] \tag{3.15}$$

and

$$BC = \frac{1}{4n} [(1) + a - b - ab - c - ac + bc + abc] \tag{3.16}$$

The  $ABC$  interaction is defined as the average difference between the  $AB$  interaction for the two different levels of  $C$ . Thus,

$$\begin{aligned} ABC &= \frac{1}{4n} \{ [abc - bc] - [ac - c] - [ab - b] + [a - (1)] \} \\ &= \frac{1}{4n} [abc - bc - ac + c - ab + b + a - (1)] \end{aligned} \tag{3.17}$$

As before, we can think of the  $ABC$  interaction as the difference in two averages. If the runs in the two averages are isolated, they define the vertices of the two tetrahedra that make up the cube in Fig. 3.6c.

**TABLE 3.4 Algebraic Signs for Calculating Effects in the 2<sup>3</sup> Design**

Treatment Combination	Factorial Effect							
	<i>I</i>	<i>A</i>	<i>B</i>	<i>AB</i>	<i>C</i>	<i>AC</i>	<i>BC</i>	<i>ABC</i>
(1)	+	−	−	+	−	+	+	−
<i>a</i>	+	+	−	−	−	−	+	+
<i>b</i>	+	−	+	−	−	+	−	+
<i>ab</i>	+	+	+	+	−	−	−	−
<i>c</i>	+	−	−	+	+	−	−	+
<i>ac</i>	+	+	−	−	+	+	−	−
<i>bc</i>	+	−	+	−	+	−	+	−
<i>abc</i>	+	+	+	+	+	+	+	+

In Equations 3.11 through 3.17, the quantities in brackets are contrasts in the treatment combinations. A table of plus and minus signs can be developed for the contrasts and is shown in Table 3.4. Signs for the main effects are determined by associating a plus with the high level and a minus with the low level. Once the signs for the main effects have been established, the signs for the interaction columns can be obtained by multiplying the appropriate preceding columns, row by row. For example, the signs in the *AB* column are the products of the *A* and *B* column signs in each row. The contrast for any effect can be obtained from this table.

Table 3.4 has several interesting properties: (1) Except for column *I*, every column has an equal number of plus and minus signs. (2) The sum of the products of the signs in any two columns is zero. (3) Column *I* multiplied by any column leaves that column unchanged. That is, *I* is an identity element. (4) The product of any two columns yields a column in the table. For example,  $A \times B = AB$ , and

$$AB \times B = AB^2 = A$$

We see that the exponents in the products are formed by using modulus 2 arithmetic. (That is, the exponent can only be zero or one; if it is greater than one, it is reduced by multiples of two until it is either zero or one.) All of these properties result from the fact that the 2<sup>3</sup> (indeed all 2<sup>*k*</sup> designs) is an **orthogonal design**.

Sums of squares for the effects are easily computed, because each effect has a corresponding single-degree-of-freedom contrast. In the 2<sup>3</sup> design with *n* replicates, the sum of squares for any effect is

$$SS = \frac{(\text{contrast})^2}{8n} \quad (3.18)$$

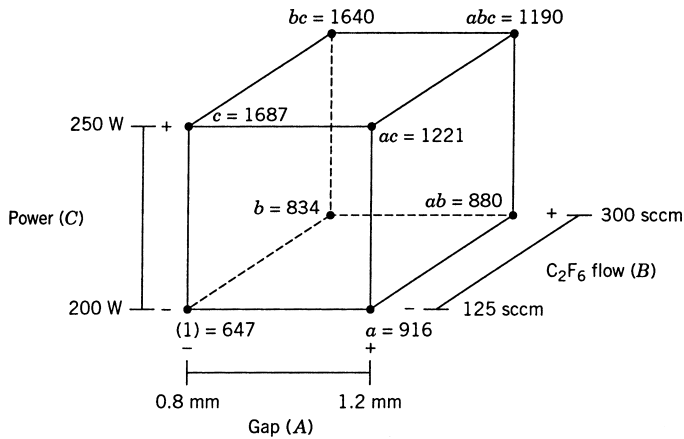
**Example 3.1 The Plasma Etch Experiment** An electrical engineer is investigating a plasma etching process used in semiconductor manufacturing. He is studying the effects of three factors—anode-cathode gap (*A*), C<sub>2</sub>F<sub>6</sub> gas flow rate (*B*), and power applied to the cathode (*C*)—on the etch rate. Each factor is run at two levels. Two replicates of a 2<sup>3</sup> factorial design are shown in Table 3.5, along with the resulting response variable values. The design is shown graphically in Fig. 3.7.

**TABLE 3.5** The 2<sup>3</sup> Factorial Design for Example 3.1

A (Gap)	B (C <sub>2</sub> F <sub>6</sub> Flow)	C (Power)	Etch Rate (Å/min)		Totals
			Replicate I	Replicate II	
-1	-1	-1	247	400	(1) = 647
1	-1	-1	470	446	a = 916
-1	1	-1	429	405	b = 834
1	1	-1	435	445	ab = 880
-1	-1	1	837	850	c = 1687
1	-1	1	551	670	ac = 1221
-1	1	1	775	865	bc = 1640
1	1	1	660	530	abc = 1190

Using the totals shown in Table 3.5 for each of the eight runs in the design, we may calculate the effect estimates as follows:

$$\begin{aligned}
 A &= \frac{1}{4n} [a - (1) + ab - b + ac - c + abc - bc] \\
 &= \frac{1}{8} [916 - 647 + 880 - 834 + 1221 - 1687 + 1190 - 1640] \\
 &= \frac{1}{8} [-601] = -75.125 \\
 B &= \frac{1}{4n} [b + ab + bc + abc - (1) - a - c - ac] \\
 &= \frac{1}{8} [834 + 880 + 1640 + 1190 - 647 - 916 - 1687 - 1221] \\
 &= \frac{1}{8} [73] = 9.125
 \end{aligned}$$



**Figure 3.7** The 2<sup>3</sup> design for Example 3.1.

$$\begin{aligned}
C &= \frac{1}{4n}[c + ac + bc + abc - (1) - a - b - ab] \\
&= \frac{1}{8}[1687 + 1221 + 1640 + 1190 - 647 - 916 - 834 - 886] \\
&= \frac{1}{8}[2461] = 307.625 \\
AB &= \frac{1}{4n}[ab - a - b + (1) + abc - bc - ac + c] \\
&= \frac{1}{8}[880 - 916 - 834 + 647 + 1190 - 1640 - 1221 + 1687] \\
&= \frac{1}{8}[-207] = -25.875 \\
AC &= \frac{1}{4n}[(1) - a + b - ab - c + ac - bc + abc] \\
&= \frac{1}{8}[647 - 916 + 834 - 880 - 1687 + 1221 - 1640 + 1190] \\
&= \frac{1}{8}[-1231] = -153.875 \\
BC &= \frac{1}{4n}[(1) + a - b - ab - c - ac + bc + abc] \\
&= \frac{1}{8}[647 + 916 - 834 - 880 - 1687 - 1221 + 1640 + 1190] \\
&= \frac{1}{8}[-229] = -28.625
\end{aligned}$$

and

$$\begin{aligned}
ABC &= \frac{1}{4n}[abc - bc - ac + c - ab + b + a - (1)] \\
&= \frac{1}{8}[1190 - 1640 - 1221 + 1687 - 880 + 834 + 916 - 647] \\
&= \frac{1}{8}[239] = 29.875
\end{aligned}$$

The largest effects are for the gap ( $A = -75.125$ ), the power ( $C = 307.625$ ), and the gap–power interaction ( $AC = -153.875$ ).

The analysis of variance may be used to confirm the magnitude of these effects. From Equation 3.18, the sums of squares are

$$\begin{aligned}
SS_A &= \frac{(-601)^2}{16} = 22,575.1 \\
SS_B &= \frac{(73)^2}{16} = 333.1 \\
SS_C &= \frac{(2461)^2}{16} = 378,532.6
\end{aligned}$$

$$SS_{AB} = \frac{(-207)^2}{16} = 2678.1$$

$$SS_{AC} = \frac{(-1231)^2}{16} = 94,710.1$$

$$SS_{BC} = \frac{(-229)^2}{16} = 3277.6$$

and

$$SS_{ABC} = \frac{(239)^2}{16} = 3570.1$$

The analysis of variance is summarized in Table 3.6. Based on the  $P$ -values shown in this table, we concluded that gap ( $A$ ) and power ( $C$ ) are statistically significant effects, as is the gap-power ( $AC$ ) interaction.

The regression model for this process, based on the reduced model from the above analysis, is

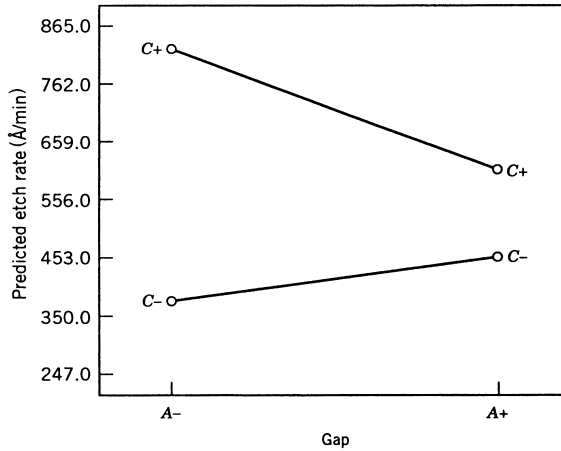
$$\begin{aligned}\hat{y} &= b_0 + b_1x_1 + b_3x_3 + b_{13}x_1x_3 \\ &= 563.438 + \left(\frac{-75.125}{2}\right)x_1 + \left(\frac{307.625}{2}\right)x_3 + \left(\frac{-153.875}{2}\right)x_1x_3 \\ &= 563.438 - 37.563x_1 + 153.813x_3 - 76.938x_1x_3\end{aligned}$$

where the variables  $x_1$  and  $x_3$  represent the design factors  $A$  and  $C$ , respectively, on the coded  $(-1, +1)$  scale. This regression model can be used to generate predicted values at each corner of the design, from which residuals can be obtained. The reader should obtain these residuals and graphically analyze them.

The factor effect estimates (and the regression model) indicate that etch rate increases as the gap ( $A$ ) decreases and as the power ( $C$ ) increases, and that these factors interact. The two-factor interaction graph, shown in Fig. 3.8, is helpful in the practical interpretation of the results. This graph was constructed by plotting average etch rate versus gap ( $A$ ) for each power setting and connecting the points for the low- and high-power settings to give the two lines shown in the figure. Inspection of the interaction graph indicates that changes

**TABLE 3.6 Analysis of Variance for Example 3.5**

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	$F_0$	$P$ -Value
$A$	22,575.1	1	22,575.1	5.64	0.0448
$B$	333.1	1	333.1	0.08	0.7802
$C$	378,532.6	1	378,532.6	94.65	$1.04 \times 10^{-5}$
$AB$	2678.1	1	2678.1	0.67	0.4369
$AC$	94,710.1	1	94,710.1	23.68	0.0012
$BC$	3277.6	1	3277.6	0.82	0.3918
$ABC$	3570.1	1	3570.1	0.89	0.3724
Error	31,995.5	8	3999.4		
Total	537,671.9	15			

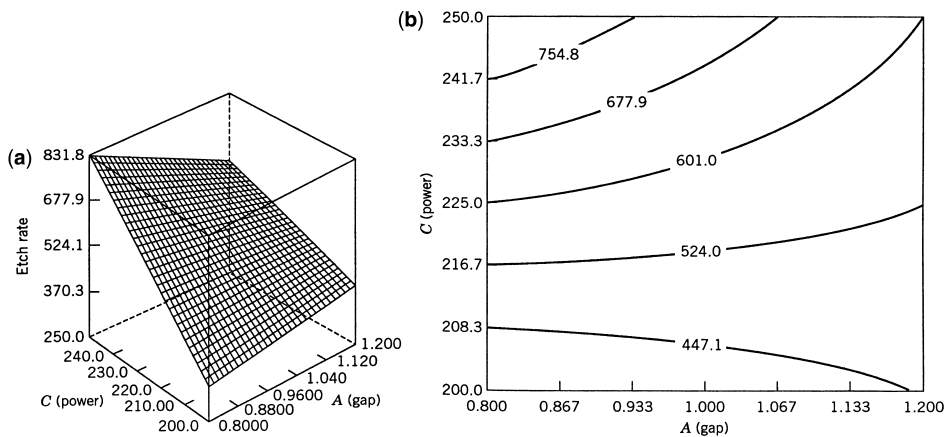


**Figure 3.8** Gap–power (AC) interaction graph, Example 3.1.

in the anode–cathode gap produces a much larger change in etch rate at the high-power setting than at the low-power setting. In addition, the impact of changing the gap from (–) to (+) differs depending on the level of power.

Figure 3.9 is a response surface plot and contour plot of etch rate as a function of gap and power setting. These plots were obtained from the fitted model. The effect of the strong interaction on this process is very clear; notice that the response surface is now a **twisted plane** (that is, the lines in the contour plot are not parallel straight lines). Thus the interaction term in the model is a form of **curvature**. In processes such as this one, the engineer is usually trying to select the appropriate factor settings (in this case gap and power) to obtain a desired etch rate. The response surface and contour plot will be most helpful for this analysis.

**Another Method for Judging the Significance of Effects** The analysis of variance is a formal way to determine which factor effects are nonzero. However, there are other methods that are useful. One method is to calculate the **standard error** of the effects and compare the magnitudes of the effects with their standard errors.



**Figure 3.9** (a) Response surface of etch rate, Example 3.1. (b) The contour plot.

The standard error of an effect is easy to find. If we assume that there are  $n$  replicates at each of the  $2^k$  runs or treatment combinations in the design, and if  $y_{i1}, y_{i2}, \dots, y_{in}$  are the observations at the  $i$ th run, then

$$s_i^2 = \frac{1}{n-1} \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2, \quad i = 1, 2, \dots, 2^k$$

is an estimate of the variance at the  $i$ th run. The  $2^k$  variance estimates can be combined to give an overall variance estimate:

$$s^2 = \frac{1}{2^k(n-1)} \sum_{i=1}^{2^k} \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2$$

This is also the variance estimate given by the error mean square in the analysis of variance. The variance of each effect estimate is

$$\begin{aligned} \text{Var}(\text{Effect}) &= \text{Var}\left(\frac{\text{Contrast}}{n2^{k-1}}\right) \\ &= \frac{1}{(n2^{k-1})^2} \text{Var}(\text{Contrast}) \end{aligned}$$

Each contrast is a linear combination of  $2^k$  treatment totals, and each total consists of  $n$  observations. Therefore,

$$\text{Var}(\text{Contrast}) = n2^k \sigma^2$$

and the variance of an effect is

$$\begin{aligned} \text{Var}(\text{Effect}) &= \frac{1}{(n2^{k-1})^2} n2^k \sigma^2 \\ &= \frac{1}{n2^{k-2}} \sigma^2 \end{aligned} \quad (3.19)$$

The estimated standard error would be found by replacing  $\sigma^2$  by its estimate  $s^2$  and taking the square root of Equation 3.19.

To illustrate this method, consider the etch rate experiment in Example 3.1. The mean square error is  $MS_E = 3999.4$ . Therefore, the standard error of each effect is (using  $s^2 = MS_E$ )

$$\begin{aligned} se(\text{Effect}) &= \sqrt{\frac{1}{n2^{k-2}} s^2} \\ &= \sqrt{\frac{1}{2(2^{3-2})} 3999.4} \\ &= 31.62 \end{aligned}$$



Two standard error limits on the effect estimates are then

$$\begin{aligned} A: & -75.125 \pm 63.24 \\ B: & 9.125 \pm 63.24 \\ C: & 307.625 \pm 63.24 \\ AB: & -25.875 \pm 63.24 \\ AC: & -153.875 \pm 63.24 \\ BC: & -28.625 \pm 63.24 \\ ABC: & 29.875 \pm 63.24 \end{aligned}$$

These intervals are approximate 95% **confidence intervals**. This analysis indicates that  $A$ ,  $C$ , and the  $AC$  interaction are important effects, because they are the only factor effect estimates for which the intervals do not include zero.

**JMP Output for Example 3.1.** Table 3.7 contains the output from JMP for the plasma etch experiment in Example 3.1. The analysis of variance for the whole model is shown near the top of the output, and the ANOVA  $F$ -statistic indicates that at least one of the seven model effects is active. Immediately below the ANOVA is a list of parameter estimates and their corresponding  $t$ -tests. Note that JMP reports the model regression coefficients as parameter estimates and not factor effect estimates. Based on the  $P$ -values, the important factors are gap, power, and the gap–power interaction. This display is followed by the effect tests, which are the corresponding tests on individual model terms from the ANOVA. Notice that the individual  $F$ -test statistics are the square of the  $t$ -test statistics. Then a table of the parameter estimates sorted in absolute magnitude is presented along with a Pareto chart reflecting their magnitude. The final part of the display is the prediction equation based on the full model fit.

The only active factors are gap, power, and the gap–power interaction. Table 3.8 presents the JMP output for the reduced model, including a plot of the residuals from the reduced model versus the fitted values of the response. This plot does not indicate any obvious problems. The bottom portion of the table presents two views of prediction profiler. In the upper display, both factors gap and power are at the middle level (0). In the bottom view, the level of gap has been changed to  $-1$ . Notice that the slope of the power curve changes in this display, reflecting the interaction between the two factors.

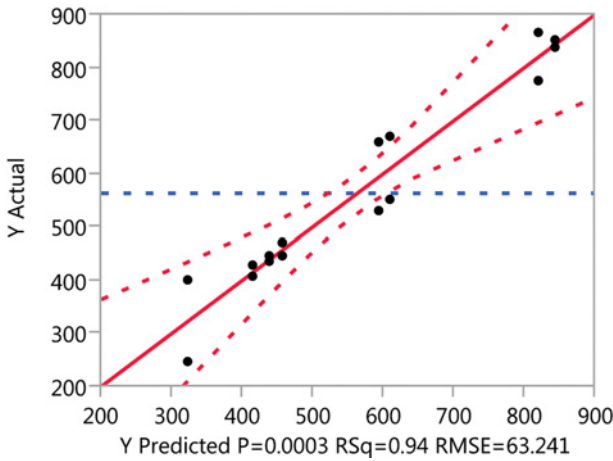
### 3.4 THE GENERAL $2^k$ DESIGN

The methods of analysis that we have presented thus far may be generalized to the case of a  $2^k$  factorial design; that is, a design with  $k$  factors, each at two levels. The statistical model for a  $2^k$  design will include  $k$  main effects,  $\binom{k}{2}$  two-factor interactions,  $\binom{k}{3}$  three-factor interactions,  $\dots$ , and one  $k$ -factor interaction. That is, for a  $2^k$  design the complete model will contain  $2^k - 1$  effects. The notation introduced earlier for treatment combinations is also used here. For example, in a  $2^5$  design  $abd$  denotes the treatment combination with factors  $A$ ,  $B$ , and  $D$  at the high level and factors  $C$  and  $E$  at the low level. The treatment

TABLE 3.7 JMP Output for Example 3.1.

Response Y

Actual by Predicted Plot



Summary of Fit

RSquare	0.940493
RSquare Adj	0.888423
Root Mean Square Error	63.24111
Mean of Response	563.4375
Observations (or Sum Wgts)	16

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F-Ratio
Model	7	505676.44	72239.5	18.0624
Error	8	31995.50	3999.4	<b>Prob &gt; F</b>
C. Total	15	537671.94		0.0003*

Parameter Estimates

Term	Estimate	Std Error	t-Ratio	Prob >  t
Intercept	563.4375	15.81028	35.64	<.0001*
Gap	-37.5625	15.81028	-2.38	0.0448*
Flow	4.5625	15.81028	0.29	0.7802
Power	153.8125	15.81028	9.73	<.0001*
Gap*Flow	-12.9375	15.81028	-0.82	0.4369
Gap*Power	-76.9375	15.81028	-4.87	0.0012*
Flow*Power	-14.3125	15.81028	-0.91	0.3918
Flow*Power*Gap	14.9375	15.81028	0.94	0.3724

TABLE 3.7 (Continued)

Effect Tests					
Source	Nparm	DF	Sum of Squares	F-Ratio	Prob > F
Gap	1	1	22575.06	5.6446	0.0448*
Flow	1	1	333.06	0.0833	0.7802
Power	1	1	378532.56	94.6465	<.0001*
Gap*Flow	1	1	2678.06	0.6696	0.4369
Gap*Power	1	1	94710.06	23.6808	0.0012*
Flow*Power	1	1	3277.56	0.8195	0.3918
Flow*Power*Gap	1	1	3570.06	0.8926	0.3724

## Sorted Parameter Estimates

Term	Estimate	Std Error	t-Ratio	t-Ratio	Prob >  t
Power	153.8125	15.81028	9.73		<.0001*
Gap*Power	-76.9375	15.81028	-4.87		0.0012*
Gap	-37.5625	15.81028	-2.38		0.0448*
Flow*Power*Gap	14.9375	15.81028	0.94		0.3724
Flow*Power	-14.3125	15.81028	-0.91		0.3918
Gap*Flow	-12.9375	15.81028	-0.82		0.4369
Flow	4.5625	15.81028	0.29		0.7802

## Prediction Expression

$$\begin{aligned}
 &563.4375 \\
 &+ -37.5625 * \text{Gap} \\
 &+ 4.5625 * \text{Flow} \\
 &+ 153.8125 * \text{Power} \\
 &+ \text{Gap} * [\text{Flow} * -12.9375] \\
 &+ \text{Gap} * [\text{Power} * -76.9375] \\
 &+ \text{Flow} * [\text{Power} * -14.3125] \\
 &+ \text{Flow} * [\text{Power} * [\text{Gap} * 14.9375]]
 \end{aligned}$$

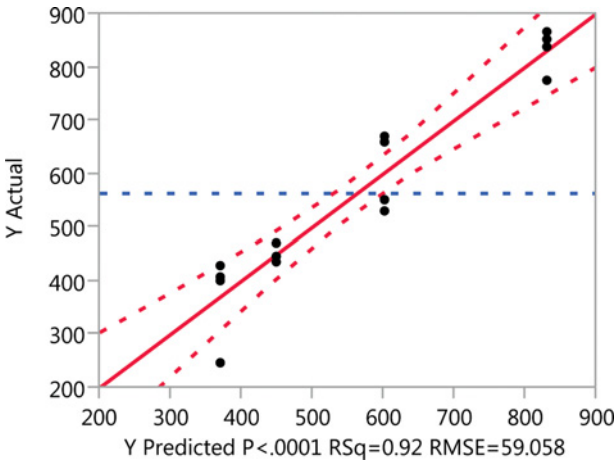
combinations may be written in standard order by introducing the factors, one at a time, with each new factor being successively combined with those that precede it. For example, the standard order for a  $2^4$  design is (1),  $a$ ,  $b$ ,  $ab$ ,  $c$ ,  $ac$ ,  $bc$ ,  $abc$ ,  $d$ ,  $ad$ ,  $bd$ ,  $abd$ ,  $cd$ ,  $acd$ ,  $bcd$ , and  $abcd$ .

To estimate an effect or to compute the sum of squares for an effect, we must first determine the contrast associated with that effect. This can always be done by using a table of plus and minus signs, such as Table 3.2 or Table 3.3. However, for large values of  $k$  this is awkward. Most modern computer programs generate the effect estimates, the regression coefficients, and the sums of squares for the analysis of variance using the method of least squares.

**TABLE 3.8 JMP Output for Example 3.1, Reduced Model**

**Response Y**

**Actual by Predicted Plot**



**Summary of Fit**

RSquare	0.922157
RSquare Adj	0.902696
Root Mean Square Error	59.05806
Mean of Response	563.4375
Observations (or Sum Wgts)	16

**Analysis of Variance**

Source	DF	Sum of Squares	Mean Square	F-Ratio
Model	3	495817.69	165273	47.3852
Error	12	41854.25	3488	<b>Prob &gt; F</b>
C. Total	15	537671.94		<.0001*

**Parameter Estimates**

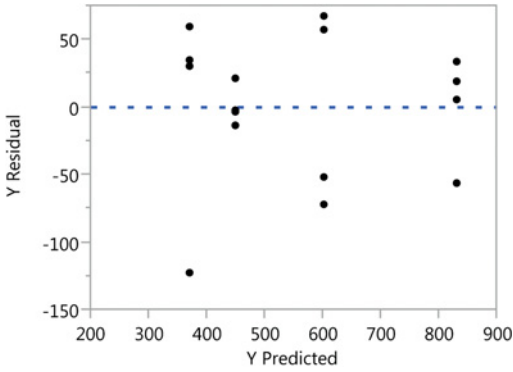
Term	Estimate	Std Error	t-Ratio	Prob >  t
Intercept	563.4375	14.76451	38.16	<.0001*
Gap	-37.5625	14.76451	-2.54	0.0257*
Power	153.8125	14.76451	10.42	<.0001*
Gap*Power	-76.9375	14.76451	-5.21	0.0002*

**Effect Tests**

Source	Nparm	DF	Sum of Squares	F-Ratio	Prob > F
Gap	1	1	22575.06	6.4725	0.0257*
Power	1	1	378532.56	108.5288	<.0001*
Gap*Power	1	1	94710.06	27.1542	0.0002*

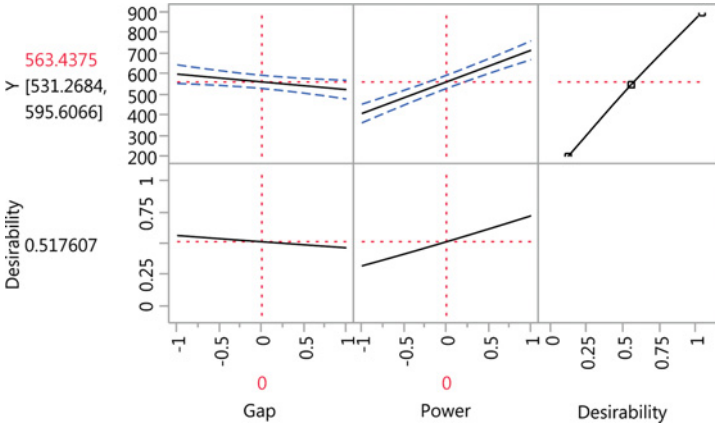
TABLE 3.8 (Continued)

Residual by Predicted Plot

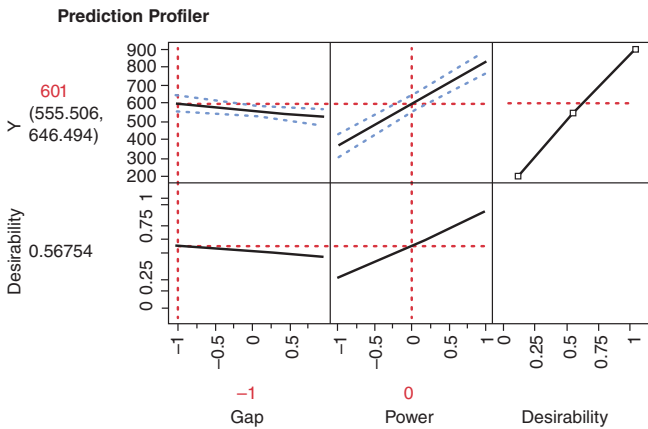


Prediction Expression

$$563.4375 + -37.5625 * \text{Gap} + 153.8125 * \text{Power} + \text{Gap} * (\text{Power} * -76.9375)$$



Prediction Profiler



### 3.5 A SINGLE REPLICATE OF THE $2^k$ DESIGN

For even a moderate number of factors, the total number of treatment combinations in a  $2^k$  factorial design is large. For example, a  $2^5$  design has 32 treatment combinations, a  $2^6$  design has 64 treatment combinations, and a  $2^{10}$  design has 1024 treatment combinations. Because resources are usually limited, the number of replicates that the experimenter can employ may be restricted. Frequently, available resources only allow a single replicate of the design to be run, unless the experimenter is willing to omit some of the original factors.

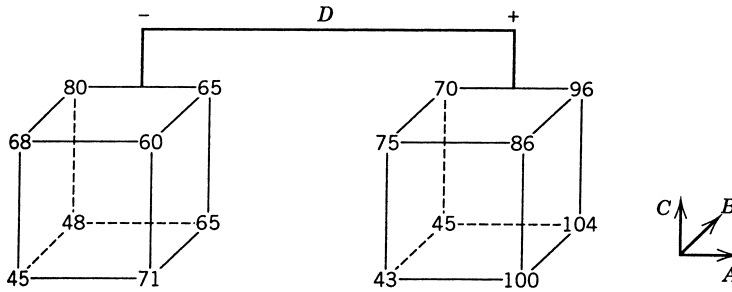
A **single replicate** of a  $2^k$  design is sometimes called an **unreplicated factorial**. With only one replicate, there is no estimate of error. One approach to the analysis of an unreplicated factorial is to assume that certain high-order interactions are negligible and combine their mean squares to estimate the error. This is an appeal to the **sparsity of effects principle**; that is, most systems are dominated by some of the main effects and low-order interactions, and most high-order interactions are negligible.

When analyzing data from unreplicated factorial designs, occasionally very high-order interactions occur. The use of an error mean square obtained by pooling high-order interactions is inappropriate in these cases. A method of analysis attributed to Daniel (1959) provides a simple and widely used way to overcome this problem. Daniel suggests constructing a normal probability plot of the effect estimates. The effects that are negligible are normally distributed, with mean zero and variance  $\sigma^2$ , and will tend to fall along a straight line on this plot, whereas significant effects will have nonzero means and will not lie along the straight line. We will illustrate this method in the following example.

**Example 3.2 A Single Replicate of the  $2^4$  Design** Montgomery (2013) describes a factorial experiment carried out in a pilot plant to study the factors thought to influence the filtration rate of a chemical product. The four factors are temperature ( $A$ ), pressure ( $B$ ), concentration of formaldehyde ( $C$ ), and stirring rate ( $D$ ). Each factor is present at two levels, and the data obtained from a single replicate of the  $2^4$  experiment are shown in Table 3.9

**TABLE 3.9 Pilot Plant Filtration Rate Experiment**

Run Number	Factor				Treatment Combination	Filtration Rate (gal/hr)
	$A$	$B$	$C$	$D$		
1	–	–	–	–	(1)	45
2	+	–	–	–	$a$	71
3	–	+	–	–	$b$	48
4	+	+	–	–	$ab$	65
5	–	–	+	–	$c$	68
6	+	–	+	–	$ac$	60
7	–	+	+	–	$bc$	80
8	+	+	+	–	$abc$	65
9	–	–	–	+	$d$	43
10	+	–	–	+	$ad$	100
11	–	+	–	+	$bd$	45
12	+	+	–	+	$abd$	104
13	–	–	+	+	$cd$	75
14	+	–	+	+	$acd$	86
15	–	+	+	+	$bcd$	70
16	+	+	+	+	$abcd$	96



**Figure 3.10** Data from the pilot plant filtration rate experiment for Example 3.2.

and Fig. 3.10. The 16 runs are made in random order. The process engineer is interested in maximizing the filtration rate. Current process conditions give filtration rates of around 75 gal/hr. The process also currently uses the concentration of formaldehyde, factor  $C$ , at the high level. The engineer would like to reduce the formaldehyde concentration as much as possible but has been unable to do so because it always results in lower filtration rates.

We will begin the analysis of this data by constructing a normal probability plot of the effect estimates. The plus and minus signs for the contrast constants for the  $2^4$  design are shown in Table 3.10. From these contrasts, we obtain the 15 factorial effect estimates shown at the foot of each column in Table 3.8. The normal probability plot of these effects is shown in Fig. 3.11. All of the effects that lie along the line are negligible, whereas the large effects are far from the line. The important effects that emerge from this analysis are the main effects of  $A$ ,  $C$ , and  $D$  and the  $AC$  and  $AD$  interactions. Table 3.11 summarizes the analysis of variance for this design. In this analysis we have pooled the nonsignificant effects to form the error term. The  $F$ -tests reveal that the effects of  $A$ ,  $C$ ,  $D$ , and the  $AC$  and  $AD$  interactions are large.

The main effects of  $A$ ,  $C$ , and  $D$  are plotted in Fig. 3.12a. All three effects are positive, and if we considered only these main effects, we would run all three factors at the high level to maximize the filtration rate. However, it is always necessary to examine any interactions that are important. Remember that main effects do not have much meaning when they are involved in significant interactions.

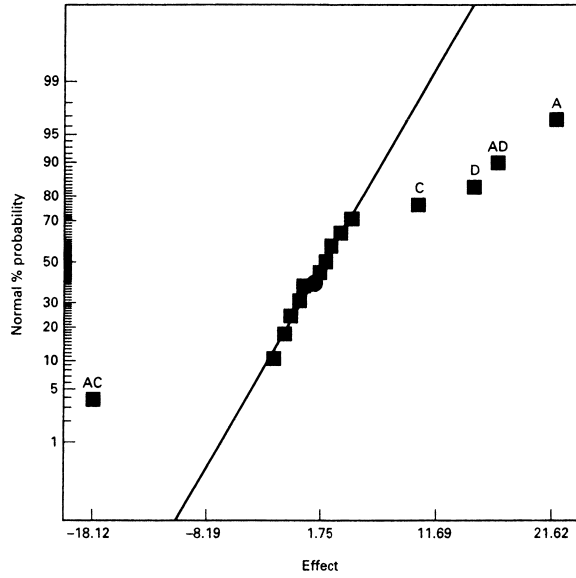
The  $AC$  and  $AD$  interactions are plotted in Fig. 3.12b. These interactions are the key to solving the problem. Note from the  $AC$  interaction that the temperature effect is very small when the concentration is at the high level and very large when the concentration is at the low level, the best results being obtained with low concentration and high temperature. The  $AD$  interaction indicates that stirring rate  $D$  has little effect at low temperature but has a large positive effect at high temperature. Therefore, the best filtration rates would appear to be obtained when  $A$  and  $D$  are at the high level and  $C$  is at the low level. This would also allow the reduction of the formaldehyde concentration to a lower level, another objective of the experimenter.

**The Half-Normal Probability Plot of Effects** An alternative to the normal probability plot of the factor effects is the **half-normal plot** (Daniel, 1976). This is a plot of the absolute value of the effect estimates against their cumulative normal probabilities. Fig. 3.13 presents the half-normal plot of the effects for the pilot plant filtration rate experiment. The straight line on the half-normal plot always passes through the origin and should also pass close to

**TABLE 3.10** Contrast Constants and Effect Estimates for the  $2^4$  Design, Example 3.2

Observations	A	B	AB	C	AC	BC	ABC	D	AD	BD	ABD	CD	ACD	BCD	ABCD
(1) = 45	-	-	+	-	+	+	-	-	+	+	-	+	-	-	+
a = 71	+	-	-	-	-	+	+	-	-	+	+	+	+	-	-
b = 48	-	+	-	-	+	-	+	-	+	-	+	+	-	+	-
b = 65	+	+	+	-	-	-	-	-	-	-	-	+	+	+	+
c = 68	-	+	+	+	-	-	+	-	+	+	-	-	+	+	-
ac = 60	+	-	-	+	+	-	-	-	-	+	+	-	-	+	+
bc = 80	-	+	-	+	-	+	-	-	+	-	+	-	+	-	+
abc = 65	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
d = 43	-	-	+	-	+	+	-	+	-	-	+	-	+	+	-
ad = 100	+	-	-	-	-	+	+	+	+	-	-	-	-	+	+
bd = 45	-	+	-	-	+	-	+	+	-	+	+	-	+	-	+
abd = 104	+	+	+	-	-	-	-	+	+	+	+	-	-	-	+
cd = 75	-	-	+	+	-	-	+	+	-	-	+	+	-	-	-
acd = 86	+	-	-	+	+	-	-	+	+	-	-	+	+	-	-
bcd = 70	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-
abcd = 96	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Effect Estimates	21.625	3.125	0.125	9.875	-18.125	2.375	1.875	14.625	16.625	-0.375	4.125	-1.125	-1.625	-2.625	1.375





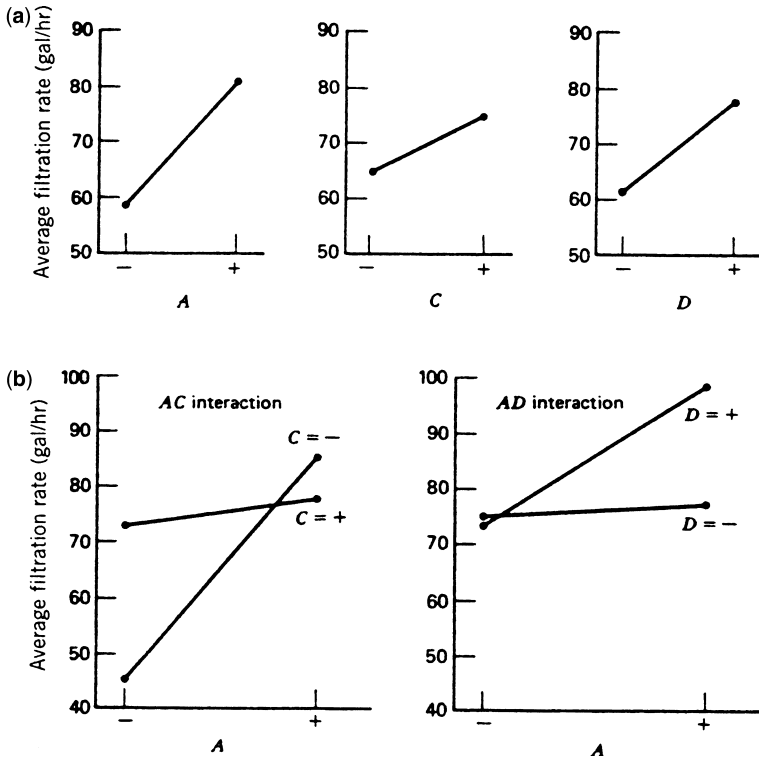
**Figure 3.11** Normal probability plot of effects for the  $2^4$  factorial in Example 3.2.

the fiftieth percentile data value. Many analysts feel that the half-normal plot is easier to interpret particularly when there are only a few effect estimates, as when the experimenter has used an eight-run design. Some software packages will construct both plots.

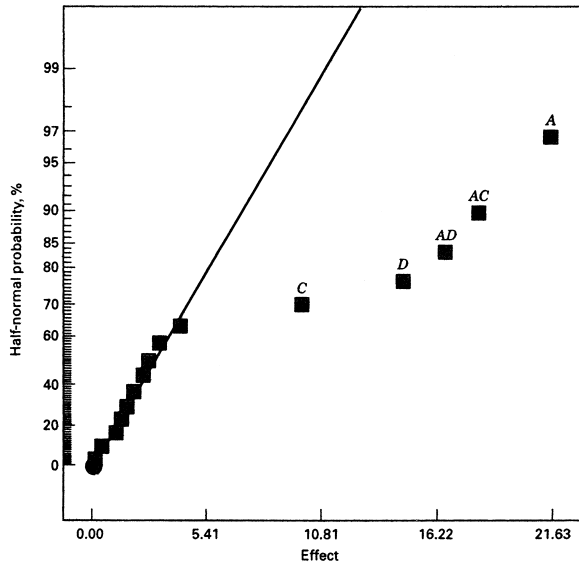
**Other Methods for Analyzing Unreplicated Factorials.** As mentioned above, the normal (or half-normal) plot of the estimated factor effects has become a standard analysis procedure for an unreplicated two-level factorial design. However, the method is not without criticism. First, it is entirely subjective, and different analysts may arrive at different conclusions from visual examination of the plot. In fact, a simple dot plot of the factor effects is very often just as reliable as the normal probability plot. Furthermore, while the normal probability plots usually work well when there are a few relatively large active effects, they may not perform well when there are several effects of small or moderate size. Because unreplicated designs are so widely used in practice, many formal analysis procedures have been proposed to overcome the subjectivity of the normal probability plot.

**TABLE 3.11** Analysis of Variance, Example 3.2

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	$F_0$	$P$ -Value
A	1870.563	1	1870.563	95.86	$1.93 \times 10^{-6}$
C	390.062	1	390.062	19.99	0.0012
D	855.562	1	855.562	43.85	0.0001
AC	1314.062	1	1314.062	67.34	$9.42 \times 10^{-6}$
AD	1105.563	1	1105.563	56.66	$2.00 \times 10^{-5}$
Error	195.125	10	19.513		
Total	5730.937	15			



**Figure 3.12** Main effect and interaction plots for Example 3.2. (a) Main effect plots, (b) Interaction plots.



**Figure 3.13** Half-normal plot of the factor effects from the  $2^4$  factorial in Example 3.2.

Hamada and Balakrishnan (1998) compared some of these methods. They found that the method proposed by Lenth (1989) has good power to detect significant effects. It is also easy to implement, and as a result it appears in several software packages for analyzing data from unreplicated factorials. We give a brief description of **Lenth's method**.

Suppose that we have  $m$  contrasts of interest, say  $c_1, c_2, \dots, c_m$ . If the design is an unreplicated  $2^k$  factorial design, these contrasts correspond to the  $m = 2^k - 1$  factor effect estimates. The basis of Lenth's method is to estimate the variance of a contrast from the smallest (in absolute value) contrast estimates. Let

$$s_0 = 1.5 \times \text{median}(|c_j|)$$

and

$$PSE = 1.5 \times \text{median}(|c_j| : |c_j| < 2.5s_0)$$

*PSE* is called the "pseudostandard error," and Lenth shows that it is a reasonable estimator of the contrast variance when there are only a few active (significant) effects. The *PSE* is used to judge the significance of contrasts. An individual contrast can be compared to the **margin of error**

$$ME = t_{0.025,d} \times PSE$$

where the degrees of freedom are defined as  $d = m/3$ . For inference on a group of contrasts, Lenth suggests using the **simultaneous margin of error**

$$SME = t_{\gamma,d} \times PSE$$

where the percentage point of the  $t$  distribution used is  $\gamma = 1 - (1 + 0.95^{1/m})/2$ .

To illustrate Lenth's method, consider the  $2^4$  experiment in Example 3.2. The calculations result in  $s_0 = 1.5 \times |-2.625| = 3.9375$  and  $2.5 \times 3.9375 = 9.84375$ , so

$$PSE = 1.5 \times |1.75| = 2.625$$

$$ME = 2.571 \times 2.625 = 6.75$$

$$SME = 5.219 \times 2.625 = 13.70$$

Now consider the effect estimates in Table 3.8. The *SME* criterion would indicate that the four largest effects (in magnitude) are significant because their effect estimates exceed *SME*. The main effect of *C* is significant according to the *ME* criterion, but not with respect to *SME*. However, because the *AC* interaction is clearly important, we would probably include *C* in the list of significant effects. Notice that in this example, Lenth's method has produced the same answer that we obtained previously from examination of the normal probability plot of effects.

Several authors [see Loughin and Nobel (1997), Hamada and Balakrishnan (1998), Lantz and Whitcomb (1998), Loughin (1998), and Edwards and Mee (2008)] have observed that Lenth's method results in values of *ME* and *SME* that are too conservative and have

relatively little power to detect significant effects. Simulation methods can be used to calibrate his procedure. Larntz and Whitcomb (1998) suggest replacing the original *ME* and *SME* multipliers with **adjusted multipliers** as follows:

Number of Contrasts	7	15	31
Original <i>ME</i>	3.764	2.571	2.218
Adjusted <i>ME</i>	2.295	2.140	2.082
Original <i>SME</i>	9.008	5.219	4.218
Adjusted <i>SME</i>	4.891	4.163	4.030

These are in close agreement with the results in Ye and Hamada (2000).

The JMP software package implements Lenth's method as part of the screening platform analysis procedure for two-level designs. In their implementation, *P*-values for each factor and interaction are computed from a "real-time" simulation. This simulation assumes that none of the factors in the experiment are significant and calculates the observed value of the Lenth statistic 10,000 times for this null model. Then *P*-values are obtained by determining where the observed Lenth statistics fall relative to the tails of these simulation-based reference distributions. These *P*-values can be used as guidance in selecting factors for the model.

Table 3.12 shows the JMP output from the screening analysis platform for the resin filtration rate experiment in Example 3.2. Notice that in addition to the Lenth statistics, the JMP output includes a half-normal plot of the effects and a "Pareto" chart of the effect (contrast) magnitudes. When the factors are entered into the model, the Lenth procedure would recommend including the same factors in the model that we identified previously.

The final JMP output for the fitted model is shown in Table 3.13. The **Prediction Profiler** at the bottom of the table has been set to the levels of the factors that maximize filtration rate. These are the same settings that we determined earlier by looking at the contour plots.

In general, the Lenth method is a clever and very useful procedure. However, we recommend using it as a **supplement** to the usual normal probability plot of effects, not as a replacement for it.

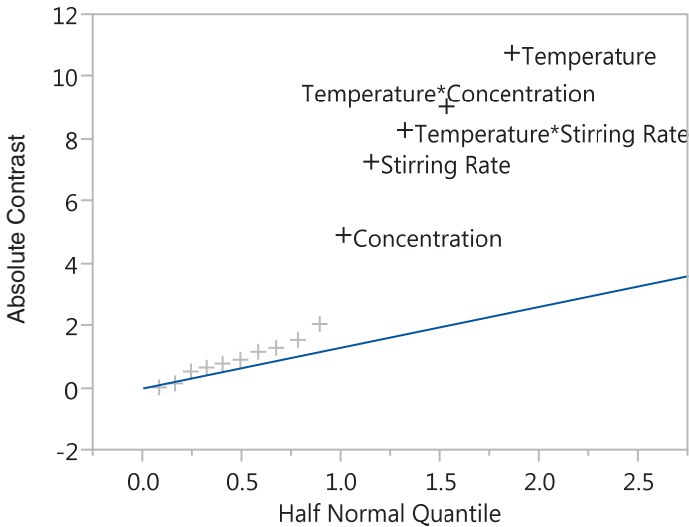
**Design Projection** Another interpretation of the data in Fig. 3.11 is possible. Because *B* (pressure) is not significant and all interactions involving *B* are negligible, we may discard *B* from the experiment so that the design becomes a  $2^3$  factorial in *A*, *C*, and *D* with two replicates. This is easily seen from examining only columns *A*, *C*, and *D* in Table 3.7 and noting that those columns form two replicates of a  $2^3$  design. The analysis of variance for the data using this simplifying assumption is summarized in Table 3.14. The conclusions that we would draw from this analysis are essentially unchanged from those of Example 3.2. Note that by projecting the single replicate of the  $2^4$  into a replicated  $2^3$ , we now have both an estimate of the *ACD* interaction and an estimate of error based on replication.

The technique of projecting an unreplicated factorial into a replicated factorial in fewer factors is very useful. In general, if we have a single replicate of a  $2^k$  design and if *h* ( $h < k$ ) factors are negligible and can be dropped, then the original data correspond to a full two-level factorial in the remaining  $k - h$  factors with  $2^h$  replicates.

**TABLE 3.12 JMP Screening Output for Example 3.2**

Screening for Filtration Rate Contrasts					
Term	Contrast	Plot of <i>t</i> -Ratio	Lenth <i>t</i> -Ratio	Individual <i>P</i> -Value	Simultaneous <i>P</i> -Value
Temperature	10.8125		8.24	0.0006*	0.0046*
Stirring Rate	7.3125		5.57	0.0022*	0.0187*
Concentration	4.9375		3.76	0.0084*	0.0777
Pressure	1.5625		1.19	0.2292	0.9581
Temperature*Stirring Rate	8.3125		6.33	0.0014*	0.0128*
Temperature*Concentration	-9.0625		-6.90	0.0010*	0.0089*
Stirring Rate*Concentration	-0.5625		-0.43	0.7009	1.0000
Temperature*Pressure	0.0625		0.05	0.9667	1.0000
Stirring Rate*Pressure	-0.1875		-0.14	0.8978	1.0000
Concentration*Pressure	1.1875		0.90	0.3553	0.9992
Temperature*Stirring Rate*Concentration	-0.8125		-0.62	0.5804	1.0000
Temperature*Stirring Rate*Pressure	2.0625		1.57	0.1242	0.7607
Temperature*Concentration*Pressure	0.9375		0.71	0.4592	1.0000
Stirring Rate*Concentration*Pressure	-1.3125		-1.00	0.3063	0.9945
Temperature*Stirring Rate*Concentration*Pressure	0.6875		0.52	0.6424	1.0000

**Half-Normal Plot**

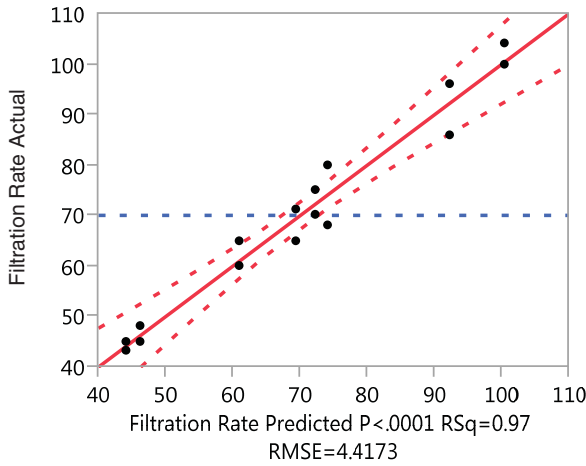


Lenth PSE = 1.3125

*P*-values derived from a simulation of 10000 Lenth *t*-ratios.

**TABLE 3.13 Final JMP Model Fit for Example 3.2**

**Response Filtration Rate  
Actual by Predicted Plot**



**Summary of Fit**

RSquare	0.965952
RSquare Adj	0.948929
Root Mean Square Error	4.417296
Mean of Response	70.0625
Observations (or Sum Wgts)	16

**Analysis of Variance**

Source	DF	Sum of Squares	Mean Square	F-Ratio
Model	5	5535.8125	1107.16	56.7412
Error	10	195.1250	19.51	<b>Prob &gt; F</b>
C. Total	15	5730.9375		<.0001*

**Lack Of Fit**

Source	DF	Sum of Squares	Mean Square	F-Ratio
Lack Of Fit	2	15.62500	7.8125	0.3482
Pure Error	8	179.50000	22.4375	<b>Prob &gt; F</b>
Total Error	10	195.12500		0.7162
				<b>Max RSq</b>

**Parameter Estimates**

Term	Estimate	Std Error	t-Ratio	Prob >  t
Intercept	70.0625	1.104324	63.44	<.0001*
Temperature	10.8125	1.104324	9.79	<.0001*
Stirring Rate	7.3125	1.104324	6.62	<.0001*
Concentration	4.9375	1.104324	4.47	0.0012*
Temperature*Stirring Rate	8.3125	1.104324	7.53	<.0001*
Temperature*Concentration	-9.0625	1.104324	-8.21	<.0001*

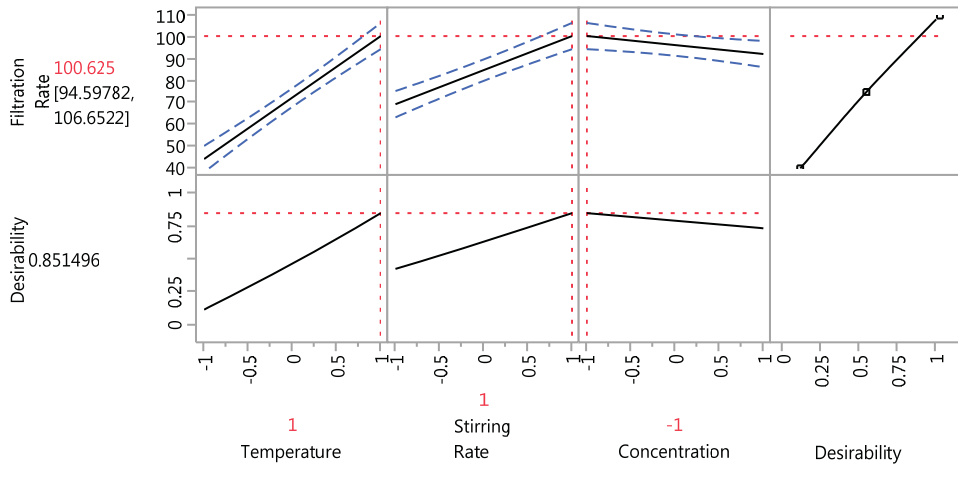
TABLE 3.13 (Continued)

Effect Tests					
Source	Nparm	DF	Sum of Squares	F-Ratio	Prob > F
Temperature	1	1	1870.5625	95.8648	<.0001*
Stirring Rate	1	1	855.5625	43.8469	<.0001*
Concentration	1	1	390.0625	19.9904	0.0012*
Temperature*Stirring Rate	1	1	1105.5625	56.6592	<.0001*
Temperature*Concentration	1	1	1314.0625	67.3447	<.0001*

Sorted Parameter Estimates

Term	Estimate	Std Error	t-Ratio	t-Ratio	Prob >  t
Temperature	10.8125	1.104324	9.79		<.0001*
Temperature*Concentration	-9.0625	1.104324	-8.21		<.0001*
Temperature*Stirring Rate	8.3125	1.104324	7.53		<.0001*
Stirring Rate	7.3125	1.104324	6.62		<.0001*
Concentration	4.9375	1.104324	4.47		0.0012*

Prediction Profiler



**The Regression Model and Diagnostic Checking** The usual diagnostic checks should be applied to the residuals from a 2<sup>k</sup> design. Our analysis indicates that the only significant effect are A = 21.625, C = 9.875, D = 14.625, AC = -18.125, and AD = 16.625. If this is true, then the predicted values of the filtration rate over the region spanned by the design are given by the regression model

$$\hat{y} = 70.06 + \left(\frac{21.625}{2}\right)x_1 + \left(\frac{9.875}{2}\right)x_3 + \left(\frac{14.625}{2}\right)x_4 - \left(\frac{18.125}{2}\right)x_1x_3 + \left(\frac{16.625}{2}\right)x_1x_4$$

**TABLE 3.14** Analysis of Variance for the Pilot Plant Filtration Rate Experiment in *A*, *C*, and *D*

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	$F_0$	$P$ -Value
<i>A</i>	1870.56	1	1870.56	83.36	$1.67 \times 10^{-5}$
<i>C</i>	390.06	1	390.06	17.38	0.0031
<i>D</i>	855.56	1	855.56	38.13	0.0003
<i>AC</i>	1314.06	1	1314.06	58.56	0.0001
<i>AD</i>	1105.56	1	1105.56	49.27	0.0001
<i>CD</i>	5.06	1	5.06	0.23	0.6476
<i>ACD</i>	10.56	1	10.56	0.47	0.5121
Error	179.52	8	22.44		
Total	5730.94	15			

where 70.06 is the average response and the coded variables  $x_1, x_3, x_4$  take on values in the interval from  $-1$  to  $+1$ . For example, the predicted filtration rate at the run where all three factors are at the low level is

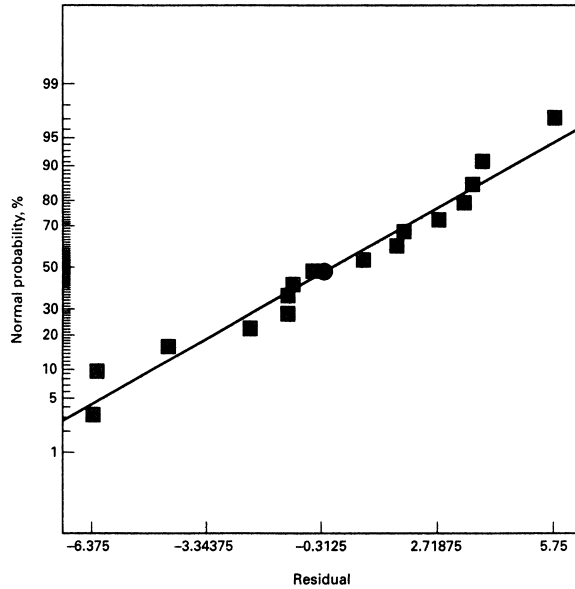
$$\begin{aligned}\hat{y} &= 70.06 + \left(\frac{21.625}{2}\right)(-1) + \left(\frac{9.875}{2}\right)(-1) + \left(\frac{14.625}{2}\right)(-1) \\ &\quad - \left(\frac{18.125}{2}\right)(-1)(-1) + \left(\frac{16.625}{2}\right)(-1)(-1) \\ &= 46.22\end{aligned}$$

Because the observed value is 45, the residual is  $e = y - \hat{y} = 45 - 46.22 = -1.22$ . The values of  $y$ ,  $\hat{y}$ , and  $e = y - \hat{y}$  for all 16 observations follow:

	$y$	$\hat{y}$	$e = y - \hat{y}$
(1)	45	46.22	-1.22
<i>a</i>	71	69.39	1.61
<i>b</i>	48	46.22	1.78
<i>ab</i>	65	69.39	-4.39
<i>c</i>	68	74.23	-6.23
<i>ac</i>	60	61.14	-1.14
<i>bc</i>	80	74.23	5.77
<i>abc</i>	65	61.14	3.86
<i>d</i>	43	44.22	-1.22
<i>ad</i>	100	100.65	3.35
<i>bd</i>	45	44.22	0.78
<i>abd</i>	104	100.65	3.35
<i>cd</i>	75	72.23	2.77
<i>acd</i>	86	92.40	-6.40
<i>bcd</i>	70	72.23	-2.23
<i>abcd</i>	96	92.40	3.60

A normal probability plot of these residuals is shown in Fig. 3.14. The points on this plot lie reasonably close to a straight line, lending support to our conclusion that *A*, *C*, *D*, *AC*,





**Figure 3.14** Normal probability plot of residuals of Example 3.2.

and  $AD$  are the only significant effects and that the underlying assumptions of the analysis are satisfied.

**Example 3.3 Duplicate Measurements of the Response** A team of engineers at a semiconductor manufacturer ran a  $2^4$  factorial design in a vertical oxidation furnace. The purpose of the experiment was to study the effects of the four factors in oxide thickness. Four wafers are stacked in the furnace, and the response variable of interest is the oxide thickness on the wafers. The four design factors are temperature ( $A$ ), time ( $B$ ), pressure ( $C$ ), and gas flow ( $D$ ). The experiment is conducted by loading four wafers into the furnace, setting the process variables to the test conditions required by the experimental design, processing the wafers, and then measuring the oxide thickness on all four wafers. Table 3.15 presents the design and the resulting thickness measurements. In this table, the four columns labeled “Thickness” contains the oxide thickness measurements on each individual wafer, and the last two columns contain the sample average and sample variance of the thickness measurements on the four wafers in each run.

The proper analysis of this experiment is to consider the individual wafer thickness measurements as **duplicate measurements**, and not as replicates. If they were actually replicates, each wafer would have been processed individually on a single run of the furnace. However, because all four wafers were processed together, they received the treatment factors (i.e., the levels of the design variables) *simultaneously*, so there is much less variability in the individual wafer thickness measurements than would have been observed if each wafer had been a genuine replicate. Therefore, the **average** of the thickness measurements is the correct response variable to consider initially.

Table 3.16 presents the effect estimates for this experiment, using the average oxide thickness  $\bar{y}$  as the response variable. Note that factors  $A$  and  $B$  and the  $AB$  interaction have large effects that together account for nearly 90% of the variability in average oxide

**TABLE 3.15 The Oxide Thickness Experiment**

Standard Order	Run Order	A	B	C	D	Thickness				$\bar{y}$	$s^2$
1	10	-1	-1	-1	-1	378	376	379	379	378	2
2	7	1	-1	-1	-1	415	416	416	417	416	0.67
3	3	-1	1	-1	-1	380	379	382	383	381	3.33
4	9	1	1	-1	-1	450	446	449	447	448	3.33
5	6	-1	-1	1	-1	375	371	373	369	372	6.67
6	2	1	-1	1	-1	391	390	388	391	390	2
7	5	-1	1	1	-1	384	385	386	385	385	0.67
8	4	1	1	1	-1	426	433	430	431	430	8.67
9	12	-1	-1	-1	1	381	381	375	383	380	12.00
10	16	1	-1	-1	1	416	420	412	412	415	14.67
11	8	-1	1	-1	1	371	372	371	370	371	0.67
12	1	1	1	-1	1	445	448	443	448	446	6
13	14	-1	-1	1	1	377	377	379	379	378	1.33
14	15	1	-1	1	1	391	391	386	400	392	34
15	11	-1	1	1	1	375	376	376	377	376	0.67
16	13	1	1	1	1	430	430	428	428	429	1.33

thickness. Figure 3.15 is a normal probability plot of the effects. From examination of this display, we conclude that factors *A*, *B*, and *C* and the *AB* and *AC* interactions are important. The analysis of variance display for this model is shown in Table 3.17.

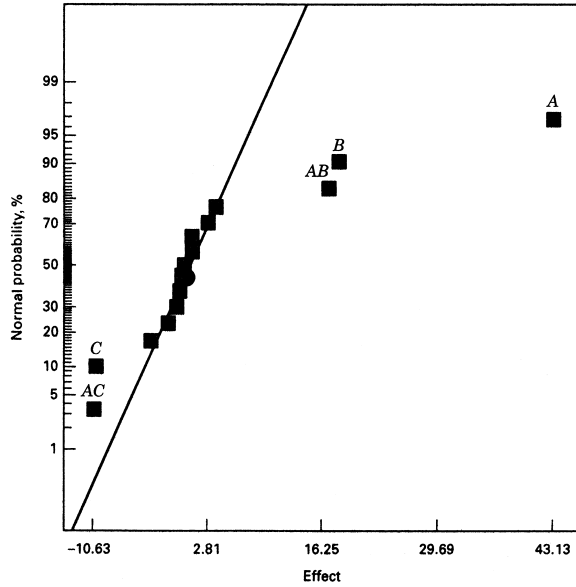
The model for predicting the average oxide thickness is

$$\hat{y} = 399.19 + 21.56x_1 + 9.06x_2 - 519x_3 + 8.44x_1x_2 - 5.31x_1x_3$$

The residual analysis for this model is satisfactory.

**TABLE 3.16 Effect Estimates for Example 3.3: Response Variable Is Average Oxide Thickness**

Model Term	Effect Estimate	Sum of Squares	Percent Contribution
<i>A</i>	43.125	7439.06	67.9339
<i>B</i>	18.125	1314.06	12.0001
<i>C</i>	-10.375	430.562	3.93192
<i>D</i>	-1.625	10.5625	0.0964573
<i>AB</i>	16.875	1139.06	10.402
<i>AC</i>	-10.625	451.563	4.12369
<i>AD</i>	1.125	5.0625	0.046231
<i>BC</i>	3.875	60.0625	0.548494
<i>BD</i>	-3.875	60.0625	0.548494
<i>CD</i>	1.125	5.0625	0.046231
<i>ABC</i>	-0.375	0.5625	0.00513678
<i>ABD</i>	2.875	33.0625	0.301929
<i>ACD</i>	-0.125	0.0625	0.000570752
<i>BCD</i>	-0.625	1.5625	0.0142688
<i>ABCD</i>	0.125	0.0625	0.000570753

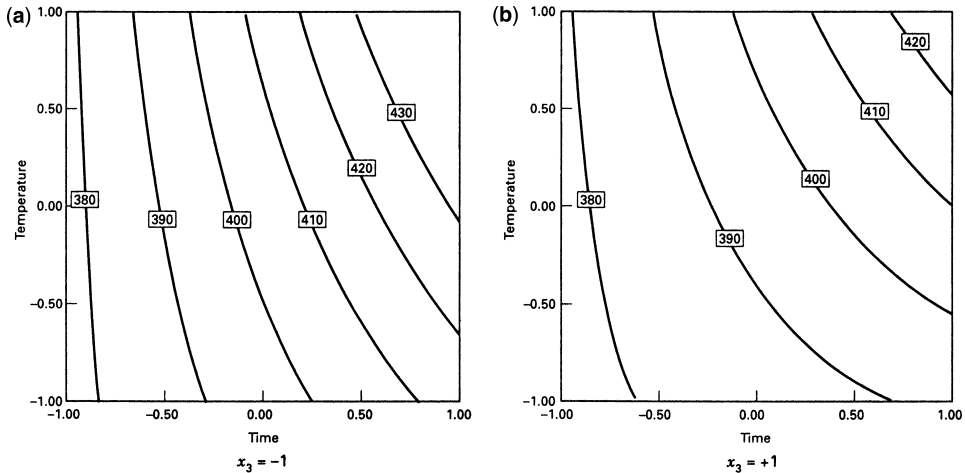


**Figure 3.15** Normal probability plot of the effects for the average oxide thickness response, Example 3.3.

**TABLE 3.17** Analysis of Variance (from Design-Expert) for the Average Oxide Thickness Response, Example 3.3

Source	Sum of Squares	DF	Mean Square	F-Value	Prob > F
<i>Model</i>	10,774.31	LC5	2154.86	122.35	<0.0000
<i>A</i>	7439.06	1	7439.06	422.37	<0.000
<i>B</i>	1314.96	1	1314.06	74.61	<0.000
<i>C</i>	430.56	1	430.56	24.45	0.0006
<i>AB</i>	1139.06	1	1139.06	64.67	<0.000
<i>AC</i>	451.56	1	451.56	25.64	0.0005
Residual	176.12	10	17.61		
Cor. Total	10,950.44	15			
Std. Dev.	4.20	<i>R-Squared</i>		0.9839	
Mean	399.19	<i>Adj. R-Squared</i>		0.9759	
C.V.	1.05	<i>Pred. R-Squared</i>		0.9588	
PRESS	450.88	<i>Adeq. Precision</i>		27.967	

Factor	Coefficient Estimate	DF	Standard Error	95% CI	
				CI Low	CI High
Intercept	399.19	1	1.05	396.85	401.53
<i>A</i> -Time	21.56	1	1.05	19.22	23.90
<i>B</i> -Temp	9.06	1	1.05	6.72	11.40
<i>C</i> -Pressure	-5.19	1	1.05	-7.53	-2.85
<i>AB</i>	8.44	1	1.05	6.10	10.78
<i>AC</i>	-5.31	1	1.05	-7.65	-2.97



**Figure 3.16** Contour plots of average oxide thickness with pressure ( $x_3$ ) held constant.

The experimenters are interested in obtaining an average oxide thickness of 400 Å, and product specifications require that the thickness must lie between 390 and 410 Å. Figure 3.16 presents two contour plots of average thickness, one with factor  $C$  (or  $x_3$ ), pressure, at the low level (i.e.,  $x_3 = -1$ ) and the other with  $C$  (or  $x_3$ ) at the high level (i.e.,  $x_3 = +1$ ). From examining these contour plots, it is obvious that there are many combinations of time and temperature (factors  $A$  and  $B$ ) that will produce acceptable results. However, if pressure is held constant at the low level, the operating window is shifted toward the left, or lower end of the time axis, indicating that lower cycle times will be required to achieve the desired oxide thickness.

It is interesting to observe the results that would be obtained if we *incorrectly* considered the individual wafer oxide thickness measurements as replicates. Table 3.18 presents a full model analysis of variance based on treating the experiment as a replicated  $2^4$  factorial. Notice that there are many significant factors in this analysis, suggesting a much more complex model than we found when using the average oxide thickness as the response. The reason for this is that the estimate of the error variance in Table 3.18 is too small ( $\hat{\sigma}^2 = 6.12$ ). The residual mean square in Table 3.18 primarily reflects the variability between wafers *within* a run and damps out the variability *between* runs. The estimate of error obtained from Table 3.17 is much larger,  $\hat{\sigma}^2 = 17.61$ , and it is primarily a measure of the between-run variability. This is the best estimate of error to use in judging the significance of process variables that are changed from run to run.

A logical question to ask is: What harm results from identifying too many factors as important?, as the incorrect analysis in Table 3.18 would certainly do. The answer is that trying to manipulate or optimize the unimportant factors would be a waste of resources, and it could result in adding unnecessary variability to *other* responses of interest.

When there are duplicate measurements on the response, there is almost always useful information about some aspect of **process variability** contained in these observations. For example, if the duplicate measurements are multiple tests by a gauge on the same experimental unit; then the duplicate measurements give some insight into **gauge capability**. If the duplicate measurements are made at different locations on an experimental unit, they

**TABLE 3.18 Analysis of Variance (from Design-Expert) of the Individual Wafer Oxide Thickness Response if Duplicates are Incorrectly Treated as Replicates**

Source	Sum of Squares	DF	Mean Square	<i>F</i> -Value	Prob > <i>F</i>
Model	43,801.75	15	2920.12	476.75	<0.0001
<i>A</i>	29,756.25	1	29,756.25	4858.16	<0.001
<i>B</i>	5256.25	1	5256.25	858.16	<0.001
<i>C</i>	1722.25	1	1722.25	281.18	<0.001
<i>D</i>	42.25	1	42.25	6.90	0.0115
<i>AB</i>	4556.25	1	4556.25	743.88	<0.001
<i>AC</i>	1806.25	1	1806.25	294.90	<0.001
<i>AD</i>	20.25	1	20.25	3.31	0.0753
<i>BC</i>	240.25	1	240.25	39.22	<0.001
<i>BD</i>	240.25	1	240.25	39.22	<0.001
<i>CD</i>	20.25	1	20.25	3.31	0.0753
<i>ABD</i>	132.25	1	132.25	21.59	<0.001
<i>ABC</i>	2.25	1	2.25	0.37	0.5473
<i>ACD</i>	0.25	1	0.25	0.941	0.8407
<i>BCD</i>	6.25	1	6.25	1.02	0.3175
<i>ABCD</i>	0.25	1	0.25	0.041	0.8407
Residual	294.00	48	6.12		
<i>Lack of Fit</i>	0.000	0			
<i>Pure Error</i>	294.00	48	6.13		
Cor. Total	44,095.75	63			

may give some information about the **uniformity** of the response variable across that unit. In our example, because we have one observation on each of four experimental units that have undergone processing together, we have some information about the **within-run variability** in the process. The information is contained in the variance of the oxide thickness measurements from the four wafers in each run. It would be of interest to determine if any of the process variables influence the within-run variability.

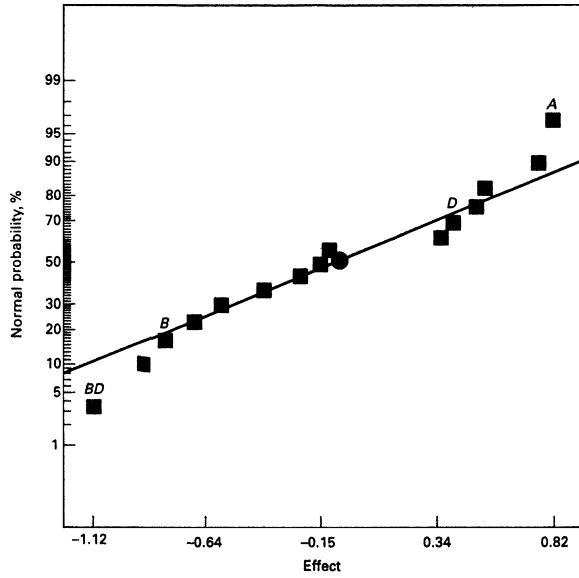
Figure 3.17 is a normal probability plot of the effect estimates obtained using  $\ln(s^2)$  as the response. As we will discuss further in Chapter 11, there is historical evidence that the log transformation is generally appropriate for modeling variability. There are not any strong individual effects, but factor *A* and the *BD* interaction are the largest. If we also include the main effects of *B* and *D* to obtain a hierarchical model, then the model for  $\ln(s^2)$  is

$$\ln(s^2) = 1.08 + 0.41x_1 - 0.40x_2 + 0.20x_4 - 0.56x_2x_4$$

The model accounts for just slightly less than half of the variability in the  $\ln(s^2)$  response, which is certainly not spectacular as empirical models go; but it is often difficult to obtain exceptionally good models of variances.

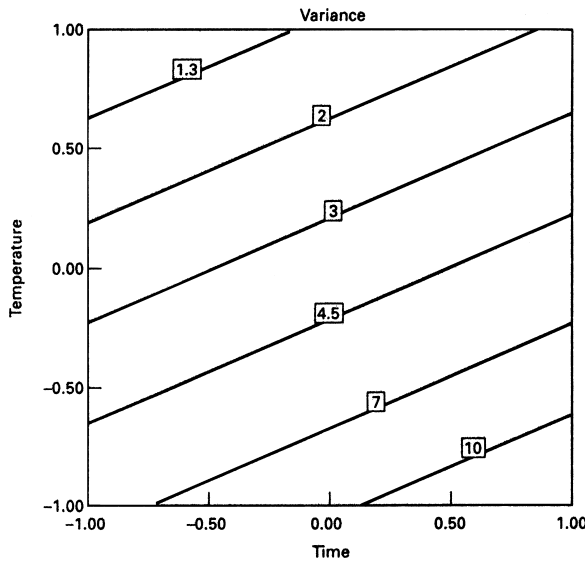
Figure 3.18 is a contour plot of the predicted variance (not the log of the predicted variance) with pressure  $x_3$  at the low level (recall that this minimizes the cycle time) and gas flow  $x_4$  at the high level. This choice of gas flow gives the lowest values of predicted variance in the region of the contour plot.

The experimenters here were interested in selecting values of the design variables that gave a mean oxide thickness within the process specifications and as close to 400 Å as

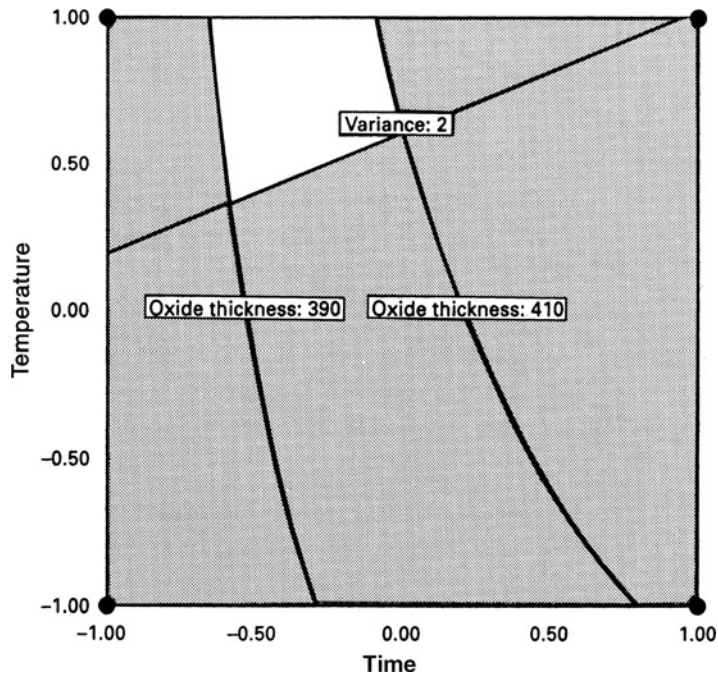


**Figure 3.17** Normal probability plot of the effects using  $\ln(s^2)$  as the response. Example 3.3.

possible, while simultaneously making the within-run variability small, say  $s^2 \leq 2$ . One possible way to find a suitable set of conditions is to overlay the contour plots in Figs. 3.16 and 3.18. The overlay plot is shown in Fig. 3.19, with the specifications on mean oxide thickness and the constraint  $s^2 \leq 2$  shown as contours. In this plot, pressure is held constant at the low level and gas flow is held constant at the high level. The open region near the upper left of the graph identifies a feasible region for the variables time and temperature.



**Figure 3.18** Contour plot of  $s^2$  (within-run variability) with pressure at the low level and gas flow at the high level.



**Figure 3.19** Overlay of the average oxide thickness and  $s^2$  responses with pressure at the low level and gas flow at the high level.

This is a simple example of using contour plots to study two responses simultaneously. We will discuss techniques for simultaneously analyzing and optimizing several responses in more detail in Chapter 7.

### 3.6 $2^k$ DESIGNS ARE OPTIMAL DESIGNS

Two-level factorial designs have many interesting and useful properties. In this section, a brief description of some of these properties is given. We showed previously in Section 3.2 that the model regression coefficients and effect estimates from a  $2^2$  design are least squares estimates. This result generalizes to the entire  $2^k$  system. For the  $2^2$  factorial design the model including the interaction term is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$$

The simplest  $2^2$  design has only  $N = 4$  runs, and it is straightforward to show that the  $\mathbf{X}'\mathbf{X}$  matrix for this design and model is

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

and the inverse of  $\mathbf{X}'\mathbf{X}$  is

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{bmatrix}$$

It turns out that the variance of any model regression coefficient is easy to find:

$$\begin{aligned} V(\hat{\beta}) &= \sigma^2(\text{diagonal element of } (\mathbf{X}'\mathbf{X})^{-1}) \\ &= \frac{\sigma^2}{4} \end{aligned}$$

All model regression coefficients have the same variance and all model regression coefficients are uncorrelated. Furthermore, there is no other four-run design on the design space bounded by  $\pm 1$  that makes the variance of the model regression coefficients smaller. In general, the variance of any model regression coefficient in a  $2^k$  design where each design point is replicated  $n$  times is  $V(\hat{\beta}) = \sigma^2/(n2^k) = \sigma^2/N$ , where  $N$  is the total number of runs in the design. This is the minimum possible variance for the regression coefficients.

For the  $2^2$  design the determinant of the  $\mathbf{X}'\mathbf{X}$  matrix is

$$|\mathbf{X}'\mathbf{X}| = 256$$

This is the maximum possible value of the determinant for a four-run design on the design space bounded by  $\pm 1$ . It turns out that the volume of the joint confidence region that contains all the model regression coefficients is inversely proportional to the square root of the determinant of  $\mathbf{X}'\mathbf{X}$ . Therefore, to make this joint confidence region as small as possible, we would want to choose a design that makes the determinant of  $\mathbf{X}'\mathbf{X}$  as large as possible. This is accomplished by choosing the  $2^2$  design.

In general, a design that minimizes the variance of the model regression coefficients is called a ***D*-optimal design**. The *D* terminology is used because these designs are found by selecting runs in the design to maximize the determinant of  $\mathbf{X}'\mathbf{X}$ . The  $2^k$  design is a *D*-optimal design for fitting the first-order model or the first-order model with interaction. Many computer software packages, including JMP, Design-Expert, and Minitab, have algorithms for finding *D*-optimal designs. These algorithms can be very useful in constructing experimental designs for many practical situations. We will make use of them in subsequent chapters.

Now consider the variance of the predicted response in the  $2^2$  design

$$V[\hat{y}(x_1, x_2)] = V(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_{12} x_1 x_2)$$

The variance of the predicted response is a function of the point in the design space where the prediction is made ( $x_1$  and  $x_2$ ) and the variance of the model regression coefficients. The



estimates of the regression coefficients are independent because the  $2^2$  design is orthogonal and they all have variance  $\sigma^2/4$ , so

$$\begin{aligned} V(\hat{y}) &= V(\hat{\beta}_0 + \hat{\beta}_1x_1 + \hat{\beta}_2x_2 + \hat{\beta}_{12}x_1x_2) \\ &= \frac{\sigma^2}{4}(1 + x_1^2 + x_2^2 + x_1^2x_2^2) \end{aligned}$$

The maximum prediction variance occurs when  $x_1 = x_2 = \pm 1$  and is equal to  $\sigma^2$ . We usually want designs that produce models that have small prediction variances. To determine how good this particular design is, we need to know the best possible value of prediction variance that we can attain. It turns out that the smallest possible value of the maximum prediction variance over the design space is  $p\sigma^2/N$ , where  $p$  is the number of model parameters and  $N$  is the number of runs in the design. The  $2^2$  design has  $N = 4$  runs and the model has  $p = 4$  parameters, so the model that we fit to the data from this experiment minimizes the maximum prediction variance over the design region. A design that has this property is called a **G-optimal design**. In general,  $2^k$  designs are  $G$ -optimal designs for fitting the first-order model or the first-order model with interaction.

We can also evaluate the prediction variance at any point or points of interest in the design space. For example, when we are at the center of the design where  $x_1 = x_2 = 0$ , the prediction variance is

$$V[\hat{y}(x = 0, x_2 = 0)] = \frac{\sigma^2}{4}$$

and if we are interested in prediction at the point where  $x_1 = 1$  and  $x_2 = 0$ , the prediction variance is

$$V[\hat{y}(x = 1, x_2 = 0)] = \frac{\sigma^2}{2}$$

An alternative to evaluating the prediction variance at a lot of different points in the design space is to consider the **average prediction variance** over the design space. One way to calculate this average prediction variance is

$$I = \frac{1}{A} \int_{-1}^1 \int_{-1}^1 V[\hat{y}(x_1, x_2)] dx_1 dx_2$$

where  $A$  is the area (in general the volume) of the design space. In computing the average variance we are summing up the variance over every point in the design space and dividing by the area or volume of the region.

Usually  $I$  is called the **average prediction variance** or the **integrated variance criteria**. Other terminology is the IV or  $Q$  criteria. For a  $2^2$  design the area of the

design region is  $A = 4$ , and

$$\begin{aligned} I &= \frac{1}{A} \int_{-1}^1 \int_{-1}^1 V[\hat{y}(x_1, x_2)] dx_1 dx_2 \\ &= \frac{1}{4} \int_{-1}^1 \int_{-1}^1 \sigma^2 \frac{1}{4} (1 + x_1^2 + x_2^2 + x_1^2 x_2^2) dx_1 dx_2 \\ &= \frac{4\sigma^2}{9} \end{aligned}$$

It turns out that this is the smallest possible value of the average prediction variance that can be obtained from a four-run design used to fit a first-order model with interaction on this design space. A design with this property is called an ***I*-optimal design**. In general,  $2^k$  designs are *I*-optimal designs for fitting the first-order model or the first-order model with interaction. The JMP software will construct *I*-optimal designs. This can be very useful in constructing designs when response prediction is the goal of the experiment.

It is also possible to display the prediction variance over the design space graphically. Figure 3.20 is output from JMP illustrating three different displays of the prediction variance from a  $2^2$  design. The first graph is the **prediction variance profiler**, which plots the **unscaled prediction variance**

$$UPV = \frac{V[\hat{y}(x_1, x_2)]}{\sigma^2}$$

against the levels of each design factor. The “crosshairs” on the graphs are adjustable, so that the unscaled prediction variance can be displayed at any desired combination of the variables  $x_1$  and  $x_2$ . Here, the values chosen are  $x_1 = -1$  and  $x_2 = +1$ , for which the unscaled prediction variance is

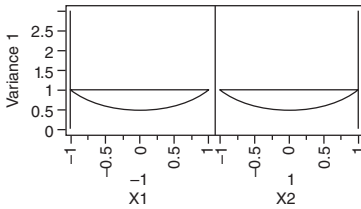
$$\begin{aligned} UPV &= \frac{V[\hat{y}(x_1, x_2)]}{\sigma^2} \\ &= \frac{\frac{\sigma^2}{4} (1 + x_1^2 + x_2^2 + x_1^2 x_2^2)}{\sigma^2} \\ &= \frac{\frac{\sigma^2}{4} (4)}{\sigma^2} \\ &= 1 \end{aligned}$$

The second graph is a **fraction of design space (FDS) plot**, which shows the unscaled prediction variance on the vertical scale and the fraction of design space on the horizontal scale. This graph also has an adjustable crosshair that is shown at the 50 percent point on the fraction of design space scale. The crosshairs indicate that the unscaled prediction variance will be at most  $0.425 \sigma^2$  (remember that the unscaled prediction variance divides by  $\sigma^2$ , that’s why the point on the vertical scale is 0.425) over a region that covers 50 percent of

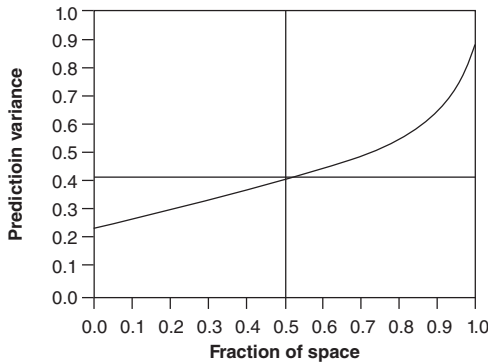
Custom Design  
Design

Run	X1	X2
1	1	1
2	1	-1
3	-1	-1
4	-1	1

Prediction Variance Profile



Fraction of Design Space Plot



Prediction Variance Surface

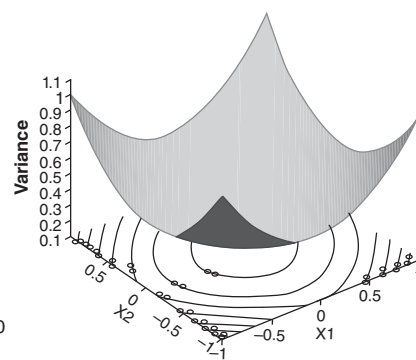


Figure 3.20 JMP prediction variance output for a  $2^2$  design.

the design region. Therefore, an FDS plot gives a simple and easy way to interpret how the prediction variance is distributed throughout the design region. An ideal FDS plot would be flat with a small value of the unscaled prediction variance. FDS plots are an ideal way to compare designs in terms of their potential prediction performance.

The final display in the JMP output is a surface plot of the unscaled prediction variance. The contours of constant prediction variance for the  $2^2$  are circular; that is, all points in the design space that are at the same distance from the center of the design have the same prediction variance. This is the property of **rotatability** that we discussed briefly in the example of fitting a second-order model using a central composite design in Section 2.8. All orthogonal  $2^k$  factorial designs have this property.

We often use optimal design algorithms to construct designs for situations where no standard design easily fits the experimental situation. For example, suppose that we are studying three continuous factors and plan to fit a first-order model with the two-factor interactions. We would like to replicate the design, which would require  $N = 16$  runs, but the experimental budget will only allow for 12 runs. Obviously we will conduct all of the runs associated with the full  $2^3$ , but where should we put the additional four runs? A  $D$ -optimal design algorithm can be used to construct the desired 12 run design.

**TABLE 3.19** A 12-Run *D*-Optimal Design from JMP

Run	$x_1$	$x_2$	$x_3$
1	-1	-1	-1
2	1	-1	-1
3	1	-1	-1
4	-1	1	-1
5	-1	1	-1
6	1	1	-1
7	-1	-1	1
8	-1	-1	1
9	1	-1	1
10	-1	1	1
11	1	1	1
12	1	1	1

Table 3.19 presents the *D*-optimal design from JMP. Notice that the four additional runs are allocated to opposite corners in opposite faces of the three-dimensional cube. The resulting 12-run design is not orthogonal. It can be shown that the correlation between any main effect regression coefficient and any two-factor interaction coefficient is 0.333. If an orthogonal 12-run design was available, the standard error of any regression coefficient would be  $\sigma/\sqrt{12} = 0.289\sigma$ . However, because the design is not orthogonal the standard error of any regression is increased to  $0.306\sigma$ . This design is not exactly *D*-optimal; it is only approximately 95% ***D*-efficient**. The *D*-efficiency is calculated by comparing the determinant of  $\mathbf{X}'\mathbf{X}$  for this design to the determinant of  $\mathbf{X}'\mathbf{X}$  for a 12-run orthogonal design. The 12-run design is not *G*-optimal either; it is approximately 86.7% ***G*-efficient**. The *G*-efficiency is computed by finding the maximum value of the prediction variance over the design space and comparing it to the theoretical minimum possible value of the maximum prediction variance  $p\sigma^2/N = 7\sigma^2/12 = 0.583\sigma^2$ .

### 3.7 THE ADDITION OF CENTER POINTS TO THE $2^k$ DESIGN

A potential concern in the use of two-level factorial designs is the assumption of *linearity* in the factor effects. Of course, perfect linearity is unnecessary, and the  $2^k$  system will work quite well even when the linearity assumption holds only very approximately. In fact, we have noted that if **interaction terms** are added to a main effects or first-order model, resulting in

$$y = \beta_0 + \sum_{j=1}^k \beta_j x_j + \sum_{i < j=2}^k \beta_{ij} x_i x_j + \varepsilon \quad (3.20)$$

then we have a model capable of representing some curvature in the response function. This curvature, of course, results from the twisting of the plane induced by the interaction terms  $\beta_{ij} x_i x_j$ .

There are going to be situations where the curvature in the response function is not adequately modeled by Equation 3.20. In such cases, a logical model to consider is the complete **second-order response surface model**:

$$y = \beta_0 + \sum_{j=1}^k \beta_j x_j + \sum_{i < j=2}^k \sum \beta_{ij} x_i x_j + \sum_{j=1}^k \beta_{jj} x_j^2 + \varepsilon \quad (3.21)$$

In running a two-level factorial experiment, we usually anticipate fitting the first-order model with interactions in Equation 3.20, but we should be alert to the possibility that the second-order model in Equation 3.21 is really more appropriate. There is a method of replicating certain points in a  $2^k$  factorial that will provide protection against curvature from second-order effects as well as allow an independent estimate of error to be obtained. The method consists of adding **center points** to the  $2^k$  design. These consist of  $n$  replicates run at the points  $x_i = 0$  ( $i = 1, 2, \dots, k$ ). One important reason for adding the replicate runs at the design center is that center points do not affect the usual effect estimates in a  $2^k$  design. When we add center points, we assume that the  $k$  factors are **quantitative**.

To illustrate the approach, consider a  $2^2$  design with one observation at each of the factorial points  $(-, -)$ ,  $(+, -)$ ,  $(-, +)$ , and  $(+, +)$  and  $n_C$  observations at the center point  $(0, 0)$ . Figure 3.21 illustrates the situation. Let  $\bar{y}_F$  be the average of the four runs at the four factorial points, and let  $\bar{y}_C$  be the average of the  $n_C$  runs at the center point. If the difference  $\bar{y}_F - \bar{y}_C$  is small, then the center points lie on or near the plane passing through the factorial points, and there is no quadratic curvature. On the other hand, if  $\bar{y}_F - \bar{y}_C$  is large, then quadratic curvature is present. A single-degree-of-freedom **sum of squares for**

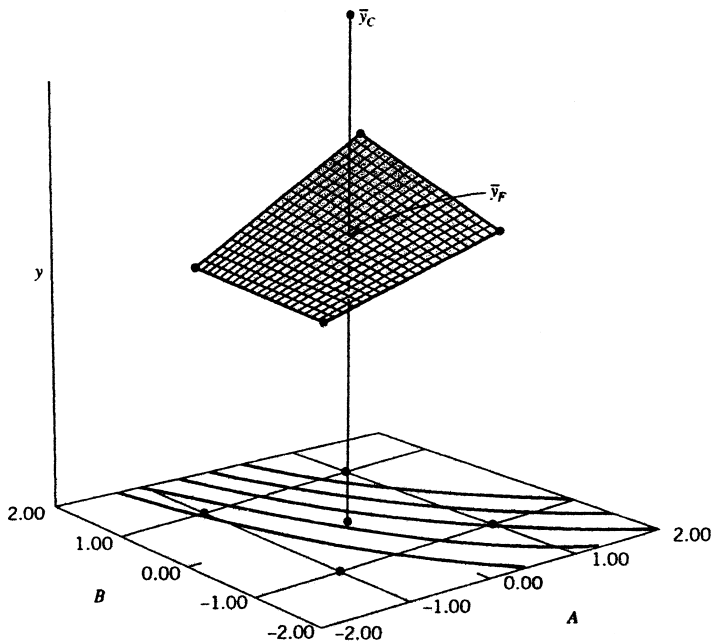
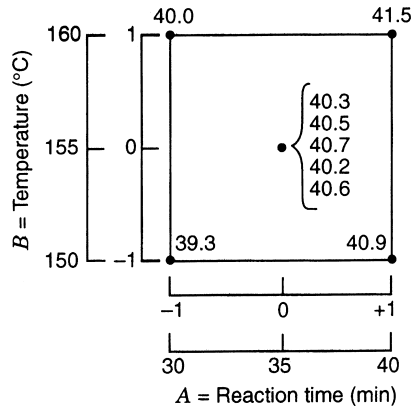


Figure 3.21 A  $2^2$  design with center points.



**Figure 3.22** The  $2^2$  design with five center points for Example 3.4.

**pure quadratic curvature** is given by

$$SS_{\text{pure quadratic}} = \frac{n_F n_C (\bar{y}_F - \bar{y}_C)^2}{n_F + n_C} \quad (3.22)$$

where, in general,  $n_F$  is the number of factorial design points. This quantity may be compared with the error mean square to test for pure quadratic curvature. More specifically, when points are added to the center of the  $2^k$  design, then the test for curvature based on Equation 3.22 actually tests the hypothesis

$$H_0: \sum_{j=1}^k \beta_{jj} = 0$$

$$H_1: \sum_{j=1}^k \beta_{jj} \neq 0$$

Furthermore, if the factorial points in the design are unreplicated, one may use the  $n_C$  center points to construct an estimate of error with  $n_C - 1$  degrees of freedom.

**Example 3.4 Testing for Curvature** A chemical engineer is studying the yield of a process. There are two variables of interest: reaction time and reaction temperature. Because she is uncertain about the assumption of linearity over the region of exploration, the engineer decides to conduct a  $2^2$  design (with a single replicate of each factorial run) augmented with five center points. The design and the yield data are shown in Fig. 3.22.

Table 3.20 summarizes the analysis of variance for this experiment. The mean square error is calculated from the center points as follows:

$$MS_E = \frac{SS_E}{n_C - 1}$$

$$= \frac{\sum_{\text{center points}} (y_i - \bar{y}_C)^2}{n_C - 1}$$

$$\begin{aligned}
&= \frac{\sum_{i=1}^5 (y_i - 40.46)^2}{4} \\
&= \frac{0.1720}{4} \\
&= 0.0430
\end{aligned}$$

The average of the points in the factorial portion of the design is  $\bar{y}_F = 40.425$ , and the average of the points at the center is  $\bar{y}_C = 40.46$ . The difference  $\bar{y}_F - \bar{y}_C = 40.425 - 40.46 = -0.035$  appears to be small. The curvature sum of squares in the analysis of variance table is computed from Equation 3.22 as follows:

$$\begin{aligned}
SS_{\text{Curvature}} &= \frac{n_F n_C (\bar{y}_F - \bar{y}_C)^2}{n_F + n_C} \\
&= \frac{(4)(5)(-0.035)^2}{4 + 5} \\
&= 0.0027
\end{aligned}$$

The analysis of variance indicates that both factors exhibit significant main effects, that there is no interaction, and that there is no evidence of curvature in the response over the region of exploration. That is, the null hypothesis  $H_0: \sum_{j=1}^2 \beta_{jj} = 0$  cannot be rejected.

In Example 3.4, we concluded that there was no indication of quadratic effects; that is, a first-order model with interaction

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$$

is appropriate (although we probably don't need the interaction term). There will be situations where the quadratic terms will be required. That is, we will have to assume a second-order model such as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon$$

Unfortunately, we cannot estimate all of the unknown parameters (the  $\beta$ 's) in this model, because there are six parameters to estimate, and the  $2^2$  design plus center points in Fig. 3.22 only has five distinct locations in the design space.

**TABLE 3.20 Analysis of Variance for Example 3.4**

Source of Variability	Sum of Squares	Degrees of Freedom	Mean Square	$F_0$	$P$ -Value
<i>A</i> (time)	2.4025	1	2.4025	55.87	0.0017
<i>B</i> (temperature)	0.4225	1	0.4225	9.83	0.0350
<i>AB</i>	0.0025	1	0.0025	0.06	0.8185
Curvature	0.0027	1	0.0027	0.06	0.8185
Error	0.1720	4	0.0430		
Total	3.0022	8			

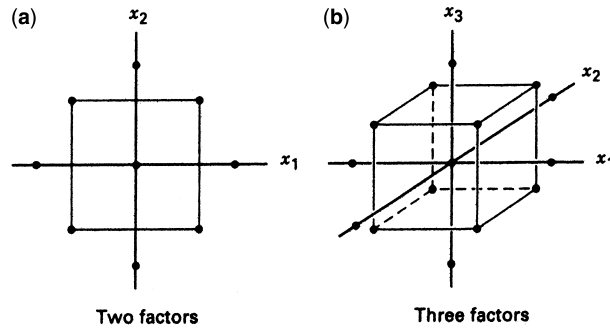


Figure 3.23 Central composite designs.

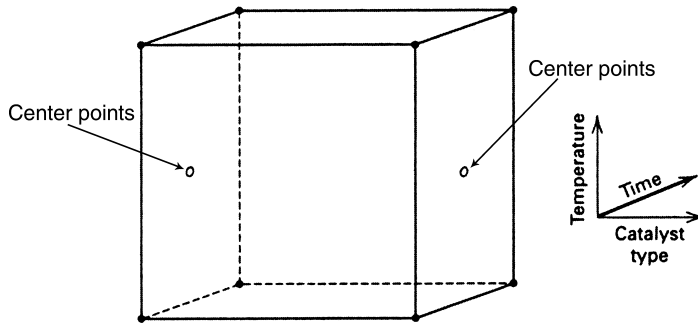
A simple and highly effective solution to this problem is to augment the  $2^2$  design with four **axial runs**, as shown in Fig. 3.23a. The resulting design is the **central composite design**. Figure 3.23b shows a central composite design for  $k = 3$  factors. This design has  $14 + n_C$  runs (usually  $3 \leq n_C \leq 5$ ), and is a very efficient design for fitting the 10-parameter second-order model in  $k = 3$  factors.

As we noted in Chapter 2, central composite designs are used extensively for fitting second-order response surface models. These designs will be illustrated and discussed in more detail in Chapters 6 and 8.

We conclude this section with a few additional useful suggestions and observations concerning the use of center points.

1. When a factorial experiment is conducted in an ongoing process, consider using the current operating conditions (or recipe) as the center point in the design. This often assures the operating personnel that at least some of the runs in the experiment are going to be performed under familiar conditions, and so the results obtained (at least for these runs) are unlikely to be any worse than are typically obtained.
2. When the center point in a factorial experiment correspond to the usual operating recipe, the experimenter can use the observed responses at the center point to provide a rough check of whether anything unusual occurred during the experiment. That is, the center point responses should be very similar to responses observed historically in routine process operation. Often operating personnel will maintain a control chart for monitoring process performance. Sometimes the center point responses can be plotted directly on the control chart as a check of the manner in which the process was operating during the experiment.
3. Consider running the replicates at the center point in nonrandom order. Specifically, run one or two center points at or near the beginning of the experiment, one or two near the middle, and one or two near the end. By spreading the center points out in time, the experimenter has a rough check on the stability of the process during the experiment. For example, if a trend has occurred in the response while the experiment was performed, plotting the center point responses versus time order may reveal this.
4. Sometimes experiments must be conducted in situations where there is little or no prior information about process variability. In these cases, running two or three center points as the first few runs in the experiment can be very helpful. These runs can provide a preliminary estimate of variability. If the magnitude of the variability seems





**Figure 3.24** A  $2^3$  factorial design with one qualitative factor and center points.

reasonable, continue; on the other hand, if larger than anticipated (or reasonable!) variability is observed, stop. Often it will be very profitable to study why the variability is so large before proceeding with the rest of the experiment.

- Usually, center points are employed when all design factors are **quantitative**. However, sometimes there will be one or more qualitative or categorical variables and several quantitative ones. Center points can still be employed in these cases. To illustrate, consider an experiment with two quantitative factors, time and temperature, each at two levels, and a single qualitative factor, catalyst type, also with two levels (organic and nonorganic). Figure 3.24 shows the  $2^3$  design for these factors. Notice that the center points are placed on the opposed faces of the cube that involve the quantitative factors. In other words, the center points can be run at the high- and low-level treatment combinations of the qualitative factors, just so long as those subspaces involve only quantitative factors.

### 3.8 BLOCKING IN THE $2^k$ FACTORIAL DESIGN

There are many situations in which it is impossible to completely randomize all of the runs in the experiment. For example, we might not be able to perform all of the runs in a factorial experiment under homogeneous conditions. As examples, these conditions might be a single time period (such as one day), a single batch of homogeneous raw material, using only one operator, or the like. In other cases, it might be desirable to deliberately vary the experimental conditions to ensure that the treatments are equally effective (one might say robust) across many situations likely to be encountered in practice. For example, a chemical engineer may run a pilot plant experiment with several batches of raw material because he knows that different raw material batches of different quality grades will be used in the full-scale process. Sometimes the presence of nuisance factors prevents complete randomization.

The design technique used in these situations is called **blocking**. In this section we show how to arrange the  $2^k$  factorial design in blocks.

#### 3.8.1 Blocking in the Replicated Design

Suppose that the  $2^k$  factorial design has been replicated  $n$  times. A simple way to incorporate nonhomogeneous conditions or nuisance factors in this situation is to consider each set of these conditions or nuisance factor levels as a block and to run each replicate of the design

**TABLE 3.21 The Chemical Process Experiment in Four Blocks (RCBD)**

	Block 1	Block 2	Block 3	Block 4
	(1) = 145 a = 158 b = 135 ab = 150	(1) = 148 a = 152 b = 138 ab = 152	(1) = 147 a = 155 b = 141 ab = 146	(1) = 140 a = 152 b = 139 ab = 149
Block totals	588	590	589	580

in a separate block. The runs in each block are to be made in random order. Because each block contains all of the runs from a complete replicate of the experiment, this type of blocking arrangement is called a **randomized complete block design (RCBD)**.

**Example 3.5 The Chemical Process Revisited** Consider the chemical process experiment described in Section 3.2. Suppose that only four experimental trials could be easily be made from a single batch of raw material. Therefore, four batches of raw material will be required in order to run all four replicates. Table 3.21 shows the RCBD for this experiment, assuming that each replicate is run in a single batch of material. Within each material batch, the four runs are made in random order.

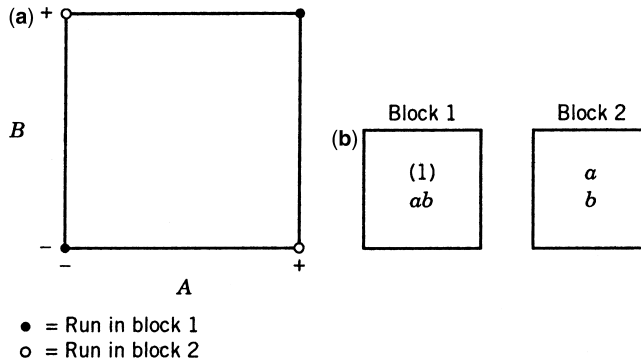
The analysis of variance for this design is shown in Table 3.22. All of the sums of squares in this table are calculated as in a standard (unblocked)  $2^k$  factorial, except the sum of squares for blocks. If we let  $B_i$  ( $i = 1, 2, 3, 4$ ) represent the four block totals (see Table 3.21), then

$$\begin{aligned}
 SS_{\text{Blocks}} &= \sum_{i=1}^4 \frac{B_i^2}{4} - \frac{y_{\dots}^2}{16} \\
 &= \frac{(588)^2 + (590)^2 + (589)^2 + (580)^2}{4} - \frac{(2347)^2}{16} \\
 &= 15.6875
 \end{aligned}$$

There are three degrees of freedom among the four blocks. Table 3.22 indicates that the conclusions from this experiment are identical to those in Section 3.2 and that the block effect is relatively small.

**TABLE 3.22 Analysis of Variance for the Chemical Process Experiment in Four Blocks, Example 3.5**

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	$F_0$	$P$ -Value
Blocks	15.6875	3	5.2292		
A	410.0625	1	410.0625	43.64	0.0001
B	138.0625	1	138.0625	14.69	0.0040
AB	3.0625	1	3.0625	0.33	0.5820
Error	84.5625	9	9.3959		
Total	651.4375	15			



**Figure 3.25** A  $2^2$  design in two blocks. (a) Geometric view. (b) Assignment of the four runs to two blocks.

### 3.8.2 Confounding in the $2^k$ Design

In many situations, it is impossible to perform a complete replicate of a factorial design in one block. **Confounding** is a design technique for arranging a complete factorial experiment in blocks, where the block size is smaller than the number of treatment combinations in one replicate. The technique causes information about certain treatment effects (usually high-order interactions) to be indistinguishable from, or **confounded** with, blocks. In this section we introduce confounding systems for the  $2^k$  factorial design. Note that even though the designs presented are **incomplete block** designs because each block does not contain all the treatments or treatment combinations, the special structure of the  $2^k$  factorial system allows a simplified method of design construction and analysis.

We consider the construction and analysis of the  $2^k$  factorial design in  $2^p$  incomplete blocks, where  $p < k$ . Consequently, these designs can be run in two blocks, four blocks, eight blocks, and so on.

**Two Blocks** Suppose that we wish to run a single replicate of the  $2^2$  design. Each of the  $2^2 = 4$  treatment combinations requires some quantity of raw material, for example, and each batch of raw material is only large enough for two treatment combinations to be tested. Thus, two batches of raw material are required. If batches of raw material are considered as blocks, then we must assign two of the four treatment combinations to each block.

Figure 3.25 shows one possible design for this problem. The geometric view, Fig. 3.25a, indicates that treatment combinations on opposing diagonals are assigned to different blocks. Notice from Fig. 3.25b. that block 1 contains the treatment combinations (1) and  $ab$  and that block 2 contains  $a$  and  $b$ . Of course the order in which the treatment combinations are run within a block is randomly determined. We would also randomly decide which block to run first. Suppose we estimate the main effect of  $A$  and  $B$  just as if no blocking had occurred. The effect estimates are

$$A = \frac{1}{2}[ab + a - b - (1)]$$

$$B = \frac{1}{2}[ab + b - a - (1)]$$

**TABLE 3.23** Table of Plus and Minus Signs for the  $2^2$  Design

Treatment Combination	Factorial Effect			
	<i>I</i>	<i>A</i>	<i>B</i>	<i>AB</i>
(1)	+	−	−	+
<i>a</i>	+	+	−	−
<i>b</i>	+	−	+	−
<i>ab</i>	+	+	+	+

Note that both of the effect estimates *A* and *B* are unaffected by blocking because in each estimate there is one plus and one minus treatment combination from each block. That is, any difference between block 1 and block 2 will cancel out.

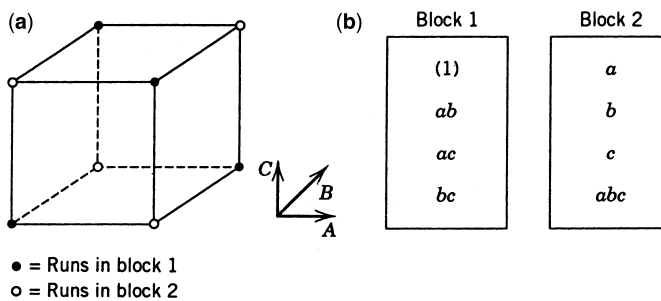
Now consider the *AB* interaction

$$AB = \frac{1}{2}[ab + (1) - a - b]$$

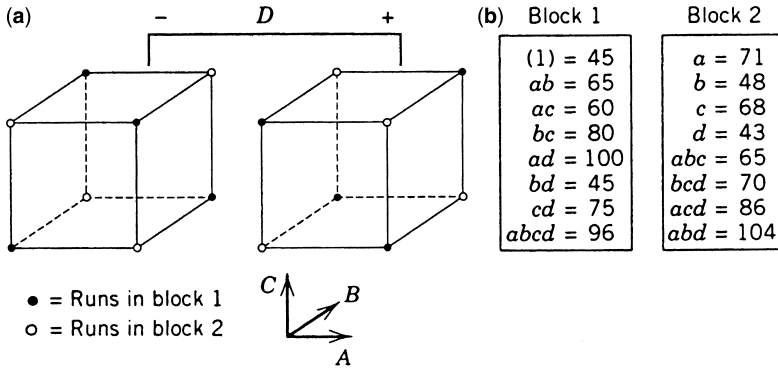
Because the two treatment combinations with the plus sign [*ab* and (1)] are in block 1 and the two with the minus sign [*a* and *b*] are in block 2, the block effect and the *AB* interaction are identical. That is, *AB* is confounded with blocks.

The reason for this is apparent from the table of plus and minus signs for the  $2^2$  design. This was originally given as Table 3.2, but for convenience it is reproduced as Table 3.23. From this table, we see that all treatment combinations that have a plus on *AB* are assigned to block 1, whereas all treatment combinations that have a minus on *AB* are assigned to block 2. This approach can be used to confound any effect (*A*, *B*, or *AB*) with blocks. For example, if (1) and *b* had been assigned to block 1 and *a* and *ab* to block 2, then the main effect *A* would have been confounded with blocks. The usual practice is to confound the highest-order interaction with blocks.

This scheme can be used to confound any  $2^k$  design in two blocks. As a second example, consider a  $2^3$  design run in two blocks. Suppose we wish to confound the three-factor interaction *ABC* with blocks. From the plus and minus signs shown in Table 3.3, we assign the treatment combinations that are minus on *ABC* to block 1 and those that are plus on *ABC* to block 2. The resulting design is shown in Fig. 3.26. Once again, we emphasize that the treatment combinations within a block are run in random order.



**Figure 3.26** A  $2^3$  design in two blocks. (a) Geometric view. (b) Assignment of the eight runs to two blocks.



**Figure 3.27** The  $2^4$  design in two blocks for Example 3.6. (a) Geometric view. (b) Assignment of the 16 runs to 2 blocks.

**Example 3.6 A  $2^4$  Factorial in Two Blocks** Consider the situation described in Example 3.2. Recall that four factors—temperature ( $A$ ), pressure ( $B$ ), concentration of formaldehyde ( $C$ ), and stirring rate ( $D$ )—are studied in a pilot plant to determine their effect on product filtration rate. Suppose now that the  $2^4 = 16$  treatment combinations cannot all be run on one day. The experimenter can make eight runs per day, so a  $2^4$  design confounded in two blocks seems appropriate. It is logical to confound the highest-order interaction  $ABCD$  with blocks. The design is shown in Fig. 3.27. Because the block totals are 566 and 555, the sum of squares for blocks is

$$SS_{\text{Blocks}} = \frac{(566)^2 + (555)^2}{8} - \frac{(1121)^2}{16} = 7.5625$$

which is identical to the sum of squares for  $ABCD$ . A normal probability plot of the remaining effects indicates that only factors  $A$ ,  $C$ ,  $D$  and the  $AC$  and  $CD$  interactions are important. Therefore, the error sum of squares is formed by pooling the remaining effects. The resulting analysis of variance is shown in Table 3.24. The practical conclusions are identical to those found in the original Example 3.2.

**TABLE 3.24 Analysis of Variance, Example 3.6**

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	$F_0$	$P$ -Value
Blocks ( $ABCD$ )	7.563	1			
$A$	1870.563	1	1870.563	89.76	$5.60 \times 10^{-6}$
$C$	390.062	1	390.062	18.72	0.0019
$D$	855.562	1	855.562	41.05	0.0001
$AC$	1314.062	1	1314.062	63.05	$2.35 \times 10^{-5}$
$AD$	1105.562	1	1105.563	53.05	$4.65 \times 10^{-5}$
Error	187.562	9	20.840		
Total	5730.937	15			

	Block 1	Block 2	Block 3	Block 4
<i>ADE</i>	-	+	-	+
<i>BCE</i>	-	-	+	+
(1)	<i>abc</i>	<i>a bc</i>	<i>b abce</i>	<i>e abcde</i>
<i>ad</i>	<i>ace</i>	<i>d abde</i>	<i>abd ae</i>	<i>ade bd</i>
<i>bc</i>	<i>cde</i>	<i>abc ce</i>	<i>c bcde</i>	<i>bce ac</i>
<i>abcd</i>	<i>bde</i>	<i>bcd acde</i>	<i>acd de</i>	<i>ab cd</i>

Figure 3.28 The  $2^5$  design in four blocks with *ADE*, *BCE*, and *ABCD* confounded.

**Four or More Blocks** It is possible to construct  $2^k$  factorial designs confounded in four blocks of  $2^k/4 = 2^{k-2}$  observations each. These designs are particularly useful in situations where the number of factors is moderately large, say  $k \geq 4$ , or block sizes are relatively small.

As an example, consider the  $2^5$  design. If each block will hold only eight runs, then four blocks must be used. The construction of this design is relatively straightforward. Select two effects to be confounded with blocks, say *ADE* and *BCE*. Then construct a table of plus and minus signs for the  $2^5$  design. If you consider only the *ADE* and *BCE* columns in such a table, there will be exactly four different sign patterns for the 32 observations in these two columns. The sign patterns are as follows:

<i>ADE</i>	<i>BCE</i>
-	-
+	-
-	+
+	+

The design would be constructed by placing all the runs that are (-, -) in block 1, those that are (+, -) in block 2, those that are (-, +) in block 3, and those that are (+, +) in block 4. The design is shown in Fig. 3.28.

With a little reflection we realize that another effect in addition to *ADE* and *BCE* must be confounded with blocks. Because there are four blocks with three degrees of freedom between them, and because *ADE* and *BCE* have only one degree of freedom each, clearly an additional effect with one degree of freedom must be confounded. This effect is the **generalized interaction** of *ADE* and *BCE*, which is defined as the product of *ADE* and *BCE* modulo 2. Thus, in our example the generalized interaction  $(ADE)(BCE) = ABCDE^2 = ABCD$  is also confounded with blocks. It is easy to verify this by referring to a table of plus and minus signs for the  $2^5$  design. Inspection of such a table would reveal that the treatment combinations are assigned to the blocks as follows:

Treatment Combinations in	Sign on <i>ADE</i>	Sign on <i>BCE</i>	Sign on <i>ABCD</i>
Block 1	-	-	+
Block 2	+	-	-
Block 3	-	+	-
Block 4	+	+	+

Notice that the product of signs of any two effects for a particular block (e.g.,  $ADE$  and  $BCE$ ) yields the sign of the other effect for that block (in this case  $ABCD$ ). Thus  $ADE$ ,  $BCE$ , and  $ABCD$  are all confounded with blocks.

The general procedure for constructing a  $2^k$  design confounded in four blocks is to choose two effects to generate the blocks, automatically confounding a third effect that is the generalized interaction of the first two. Then, the design is constructed by assigning the treatment combinations that have the sign pattern  $(-, -)$ ,  $(+, +)$ ,  $(+, -)$ ,  $(-, +)$  in the two original columns to four blocks. In selecting effects to be confounded with blocks, care must be exercised to obtain a design that does not confound effects that may be of interest. For example, in a  $2^5$  design we might choose to confound  $ABCDE$  and  $ABD$ , which automatically confounds  $CE$ , an effect that is probably of interest. A better choice is to confound  $ADE$  and  $BCE$ , which automatically confounds  $ABCD$ . It is preferable to sacrifice information on the three-factor interactions  $ADE$  and  $BCE$  instead of the two-factor interaction  $CE$ .

We can extend this procedure to even larger numbers of blocks. It is possible to construct a  $2^k$  factorial design in  $2^p$  blocks (where  $p < k$ ) of  $2^{k-p}$  treatment combinations each. To do this, select  $p$  independent effects to confound with blocks, where “independent” means that no chosen effect is the generalized interaction of the others. The plus and minus sign patterns on these  $p$  columns are used to assign the treatment combinations to the blocks. In addition, exactly  $2^p - p - 1$  other effects will be confounded with blocks, these effects being the generalized interactions of the  $p$  effects initially chosen. The choice of these  $p$  effects is critical, because the confounding pattern in the design depends on them. Table 3.25, from Montgomery (2013), presents a selection of useful designs for  $k \leq 7$  factors. To illustrate the use of this table, suppose we want to run a  $2^6$  design in  $2^3 = 8$  blocks of  $2^{6-3} = 8$  runs each. Table 3.25 indicates that we should choose  $ABEF$ ,  $ABCD$ , and  $ACE$  as the  $p = 3$  independent effects to construct the blocks. Then  $2^p - p - 1 = 2^3 - 3 - 1 = 4$  other effects are also confounded. These are the generalized interaction of these three, or

$$\begin{aligned}(ABEF)(ABCD) &= A^2B^2CDEF = CDEF \\(ABEF)(ACE) &= A^2BCE^2F = BCF \\(ABCD)(ACE) &= A^2BC^2DE = BDE \\(ABEF)(ABCD)(ACE) &= A^3B^2C^2DE^2F = ADF\end{aligned}$$

Notice that this design would confound three four-factor interactions and four three-factor interactions with blocks. This is the best possible blocking arrangement for this design.

There are several computer programs for constructing and analyzing two-level factorial designs. Most of these programs have the capability to generate the confounded versions of these designs discussed in this section. The authors of these programs have typically used Table 3.25, or a similar one, as the basis for their designs.

### 3.9 SPLIT-PLOT DESIGNS

The experimental designs we have considered so far are either **completely randomized designs** (CRDs) or **randomized blocks**, where the randomization is restricted to within different levels of nuisance factors (the blocks). In some experimental situations we may be

**TABLE 3.25 Suggested Blocking Arrangements for the  $2^k$  Factorial Design**

Number of Factors, $k$	Number of Blocks, $2^p$	Block Size, $2^{k-p}$	Effects Chosen to Generate the Blocks	Interactions Confounded with Blocks
3	2	4	$ABC$	$ABC$
	4	2	$AB, AC$	$AB, AC, BC$
4	2	8	$ABCD$	$ABCD$
	4	4	$ABC, ACD$	$ABC, ACD, BD$
	8	2	$AB, BC, CD$	$AB, BC, CD, AC, BD, AD, ABCD$
5	2	16	$ABCDE$	$ABCDE$
	4	8	$ABC, CDE$	$ABC, CDE, ABDE$
	8	4	$ABE, BCE, CDE$	$ABE, BCE, CDE, AC, ABCD, BD, ADE$
	16	2	$AB, AC, CD, DE$	All 2-factor and 4-factor interactions (15 effects)
6	2	32	$ABCDEF$	$ABCDEF$
	4	16	$ABCF, CDEF$	$ABCF, CDEF, ABDE$
	8	8	$ABEF, ABCD, ACE$	$ABEF, ABCD, ACE, BCF, BDE, CDEF, ADF$
	16	4	$ABF, ACF, BDF, DEF$	$ABE, ACF, BDF, DEF, BC, ABCD, ABDE, AD, ACDE, CE, BDF, BCDEF, ABCEF, AEF, BE$
	32	2	$AB, BC, CD, DE, EF$	All 2-factor, 4-factor, and 6-factor interactions (31 effects)
7	2	64	$ABCDEFG$	$ABCDEFG$
	4	32	$ABCFG, CDEFG$	$ABCFG, CDEFG, ABDE$
	8	16	$ABC, DEF, AFG$	$ABC, DEF, AFG, ABCDEF, DCFG, ADEG, BCDEFG$
	16	8	$ABD, EFG, CDE, ADG$	$ABCD, EFG, CDE, ADG, ABCDEFG, ABE, BCG, CDFG, ADEF, ACEG, ABFG, BCEF, BDEG, ACF, BDF$
	32	4	$ABG, BCG, CDG, DEG, EFG$	$ABG, BCG, CDG, DEG, EFG, AC, BD, DE, DF, AE, BE, ABCD, ABDE, ABEF, BCDE, BCEF, CDEF, ABCDEFG, ADG, ACDEG, ACEFG, ABDFG, ABCEG, BEG, BDEFG, CFG, ADEF, ACDF, ABCF, AFG$
	64	2	$AB, BC, CD, DE, EF, FG$	All 2-factor, 4-factor, and 6-factor interactions (63 effects)



unable to completely randomize the order of the runs because some of the factors are hard to change. To illustrate, suppose that it is hard to change the level for one of the factors in the experiment. When this happens, practitioners typically fix the level of the hard-to-change factor and run multiple combinations of the other factors at that level of the hard-to-change factor. This results in a generalization of the factorial design called a **split-plot design**.

As an example, consider a paper manufacturer who is interested in two different pulp preparation methods (the methods differ in the amount of hardwood in the pulp mixture) and two different cooking temperatures for the pulp and who wishes to study the effect of these two factors on the tensile strength of the paper. Each replicate of a factorial experiment requires four observations, and the experimenter has decided to run two replicates. Therefore, the complete experiment will require eight runs. Also, the pulp preparation method factor is hard to change in that each batch of pulp is quite large, and making up a separate batch of pulp for each run of the experiment or eight pulp batches is impractical. Therefore, he conducts the experiment as follows. A batch of pulp is produced by one of the two methods under study. Then this batch is divided into two samples, and each sample is cooked at one of the two temperatures. Then a second batch of pulp is made up using the remaining pulp preparation method. This second batch is also divided into two samples that are tested at the two temperatures. Then a third batch is made up using one of the two methods, it is divided into two samples and the samples are tested using the two temperatures. Finally, another batch of pulp is prepared using the remaining method, it is divided into two samples and both temperatures are tested. The data are shown in Table 3.26.

Initially, we might consider this to be a factorial experiment with two levels of preparation method (factor *A*) and two levels of temperature (factor *B*). If this is the case, then the order of experimentation should be completely randomized; that is, within a block, we should randomly select a treatment combination (a preparation method and a temperature) and obtain an observation, and then we should randomly select another treatment combination and obtain a second observation, and so on, until all eight treatment combinations corresponding to the two replicates have been taken. However, the experimenter did not conduct the experiment this way. He made up a batch of pulp and obtained observations for both temperatures from that batch. Because of the size of the batches, this is the only feasible way to run this experiment. A completely randomized factorial experiment would require eight batches of pulp, which is unrealistic. The split-plot design requires only four

**TABLE 3.26 Split-Plot Design for the Pulp Preparation Method–Temperature Experiment**

Whole Plots	Method	Temperature	<i>y</i>
1	–1	–1	36
1	–1	1	50
2	1	–1	25
2	1	1	30
3	–1	1	46
3	–1	–1	35
4	1	–1	20
4	1	1	27

TABLE 3.27 JMP Output for the Split-Plot Experiment in Table 3.26

Response Y									
Summary of Fit									
R-Square	0.995006								
R-Square Adj	0.991261								
Root Mean Square Error	1.274755								
Mean of Response	33.625								
Observations (or Sum Wgts)	8								
Parameter Estimates									
Term	Estimate	Std Error	DFDen	t-Ratio	Prob >  t				
Intercept	33.625	1.179248	2	28.51	0.0012				
X1	-8.125	1.179248	2	-6.89	0.0204				
X2	4.625	0.450694	2	10.26	0.0094				
X1*X2	-1.625	0.450694	2	-3.61	0.0691				
REML Variance Component Estimates									
Random Effect	Var Ratio	Var Component	Std Error	95% Lower	95% Upper	Pct of Total			
Whole Plots	2.9230769	4.75	5.6215267	-6.268192	15.768192	74.510			
Residual	1.625	1.625	1.625	0.4405132	64.184072	25.490			
Total		6.375				100.000			
-2 LogLikelihood = 25.45867972									
Fixed Effect Tests									
Source	Nparm	DF	DFDen	F-Ratio	Prob > F				
X1	1	1	2	47.4719	0.0204				
X2	1	1	2	105.3077	0.0094				
X1*X2	1	1	2	13.0000	0.0691				

batches. Obviously, the split-plot design has resulted in a gain in experimental efficiency when compared to the completely randomized design.

Split-plot designs have their origin in the agricultural sciences. Suppose that we have an agricultural field trial with two factors: irrigation method and seed variety. The irrigation method is hard to change, in that it can only be applied over a relatively large section of the test field, called a **whole plot**. The factor associated with this is therefore called a **whole plot factor**. Within the whole plot, the second factor—seed variety—is applied to smaller sections of the field, which are obtained by splitting the larger section of the field into **subplots** or **split plots**. This factor is called the **subplot factor**. In the paper manufacturing example, pulp preparation method is the whole plot factor and temperature is the subplot factor.

The pulp preparation method–temperature experiment is fairly typical of how the split-plot design is used in an industrial setting. Notice that the two factors were essentially “applied” at different times. Consequently, a split-plot design can be viewed as two experiments “combined” or superimposed on each other. One “experiment” has the whole plot factor applied to the large experimental units (or it is a factor whose levels are hard to change) and the other “experiment” has the subplot factor applied to the smaller experimental units (or it is a factor whose levels are easy to change).

The statistical model for a split-plot design can be written as

$$y_{ij} = \mu + \tau_i + \varepsilon_{WP} + \beta_j + (\tau\beta)_{ij} + \varepsilon_{SP}$$

where  $\mu$  is an overall mean,  $\tau_i$  is the whole plot factor (the pulp preparation methods),  $\varepsilon_{WP}$  is the whole plot error,  $\beta_j$  is the subplot factor (the temperature),  $(\tau\beta)_{ij}$  is the method–temperature interaction, and  $\varepsilon_{SP}$  is the subplot error. Notice that the model has two error terms,  $\varepsilon_{WP}$  and  $\varepsilon_{SP}$ , because there are two different randomizations involved; one for the whole plots, and another for the subplots. Consequently there are two variance components that represent these errors,  $\sigma_{WP}^2$  and  $\sigma_{SP}^2$ . We usually say that the split-plot design has **two different error structures**. All of the whole plot factors are tested against the whole plot error, and all of the subplot factors along with interactions between whole plot and subplot factors are tested against the subplot error. Generally we find that the subplot error variance is less than the whole plot error variance, because the subplots are usually more homogeneous than the whole plots. Because the subplot treatments are compared with greater precision, it is preferable to assign the factor we are most interested in to the subplots, if possible. However, practical or economic considerations of the hard-to-change versus easy-to-change nature of the factors may make this impossible.

JMP will construct and analyze split-plot designs. Some of the JMP output for the pulp preparation method–temperature experiment is shown in Table 3.27. In addition to an ANOVA summary for the model (see the section in Table 3.27 labeled “Fixed Effect Tests”) JMP also obtains estimates of the whole plot and subplot variance components using a maximum likelihood approach (see the section labeled “REML Variance Components Estimates”). Both factors, pulp preparation methods and temperature are significant, and there is reasonable indication of an interaction between these factors (the  $P$ -value is approximately 0.07). Also, we see that the whole plot error estimate  $\hat{\sigma}_{WP}^2 = 4.75$ , is greater than the subplot error estimate  $\hat{\sigma}_{SP}^2 = 1.625$ .

## EXERCISES

- 3.1** A router is used to cut registration notches on a printed circuit board. The vibration at the surface of the board as it is cut is considered to be a major source of dimensional variation in the notches. Two factors are thought to influence vibration: bit size ( $A$ ) and cutting speed ( $B$ ). Two bit sizes ( $\frac{1}{16}$  and  $\frac{1}{8}$  inch) and two speeds (40 and 90 rpm) are selected, and four boards are cut at each set of conditions as shown in Table E3.1. The response variable is vibration measured as the resultant vector of three accelerometers ( $x$ ,  $y$ , and  $z$ ) on each test circuit board.

TABLE E3.1 The  $2^2$  Design for Exercise 3.1

A	B	Treatment Combination	Replicate			
			I	II	III	IV
–	–	(1)	18.2	18.9	12.9	14.4
+	–	$a$	27.2	24.0	22.4	22.5
–	+	$b$	15.9	14.5	15.1	14.2
+	+	$ab$	41.0	43.9	36.3	39.9

- (a) Analyze the data from this experiment.
- (b) Construct a normal probability plot of the residuals and a plot of the residuals versus the predicted vibration level. Interpret these plots.
- (c) Draw the  $AB$  interaction plot. Interpret this plot. What levels of bit size and speed would you recommend for routine operation?
- (d) Construct a contour plot of vibration as a function of speed and bit size.
- 3.2** An engineer is interested in the effects of cutting speed ( $A$ ), tool geometry ( $B$ ), and cutting angle ( $C$ ) on the life (in hours) of a machine tool. Two levels of each factor are chosen, and two replicates of a  $2^3$  factorial design are run. The results are shown in Table E3.2.

TABLE E3.2 The  $2^3$  Design for Exercise 3.2

A	B	C	Treatment Combination	Replicate	
				I	II
–	–	–	(1)	22	31
+	–	–	$a$	32	43
–	+	–	$b$	35	34
+	+	–	$ab$	55	47
–	–	+	$c$	44	45
+	–	+	$ac$	40	37
–	+	+	$bc$	60	50
+	+	+	$abc$	39	41

- (a) Estimate the factor effects. Which effects appear to be large?
- (b) Use the analysis of variance to confirm your conclusions for part (a).

- (c) Analyze the residuals. Are there obvious problems or violations of the assumptions?
- (d) What levels of  $A$ ,  $B$ , and  $C$  would you recommend, based on the data from this experiment?
- 3.3** An industrial engineer employed by a beverage bottler is interested in the effects of two different types of 32-ounce bottles on the time to deliver 12-bottle cases of the product. The two bottle types are glass and plastic. Two workers are used to perform a task consisting of moving 40 cases of the product 50 ft on a standard type of hand truck and stacking the cases in a display. Four replicates of a  $2^2$  factorial design are performed, and the times observed are given in Table E3.3. Analyze the data and draw appropriate conclusions. Analyze the residuals and comment on model adequacy.

**TABLE E3.3** The  $2^2$  Design for Exercise 3.3

Bottle Type	Time	
	Worker 1	Worker 2
Glass	5.12	6.65
	4.98	5.49
	4.89	6.24
	5.00	5.55
Plastic	4.95	5.28
	4.27	4.75
	4.43	4.91
	4.25	4.71

- 3.4** Find the two standard error limits for the factor effects in Exercise 3.1. Does the results of this analysis agree with the conclusions from the analysis of variance?
- 3.5** Find the two standard error limits for the factor effects in Exercise 3.2. Does the result of this analysis agree with the conclusions from the analysis of variance?
- 3.6** An experiment was performed to improve the yield of a chemical process. Four factors were selected, and one replicate of a completely randomized experiment was run. The results are shown in Table E3.4.

**TABLE E3.4** The  $2^3$  Design for Exercise 3.6

Treatment Combination	Yield	Treatment Combination	Yield
(1)	90	$d$	98
$a$	64	$ad$	62
$b$	81	$bd$	87
$ab$	63	$abd$	75
$c$	77	$cd$	99
$ac$	61	$acd$	69
$bc$	88	$bcd$	87
$abc$	53	$abcd$	60

- (a) Estimate the factor effects. Construct a normal probability plot of these effects. Which effects appear large?
- (b) Prepare an analysis of variance table using the information obtained in part (a).
- (c) Plot the residuals versus the predicted yield and on normal probability paper. Does the residual analysis appear satisfactory?
- (d) Construct a contour plot of yield as a function of the important process variables.
- 3.7 Consider the design of Exercise 3.6. Suppose that four additional runs were made at the center of the region of experimentation. The response values at these center points were 94, 90, 99, and 87. Use this additional information to test for curvature in the response function. What are your conclusions?
- 3.8 An article in the *AT&T Technical Journal* (March/April 1986, Vol. 65, pp. 39–50) describes the application of two-level factorial designs to integrated circuit manufacturing. A basic processing step is to grow an epitaxial layer on polished silicon wafers. The wafers mounted on a susceptor are positioned inside a bell jar, and chemical vapors are introduced. The susceptor is rotated and heat is applied until the epitaxial layer is thick enough. An experiment was run using two factors: arsenic flow rate ( $A$ ) and deposition time ( $B$ ). Four replicates were run, and the epitaxial layer thickness (in micrometers) was measured. The data from this experiment are in Table E3.5.
- (a) Estimate the factor effects.
- (b) Conduct the analysis of variance. Which factors are important?
- (c) Analyze the residuals. Are there any residuals that should cause concern?
- (d) Discuss how you might deal with the potential outlier found in part (c).
- (e) Construct a contour plot of the layer thickness response surface as a function of flow rate and deposition time.

TABLE E3.5 The Layer Thickness Experiment from Exercise 3.8

A	B	Replicate				Level		
		I	II	III	IV	Factor	Low (–)	High (+)
–	–	14.037	16.165	13.972	13.907	A	55%	59%
+	–	13.880	13.860	14.032	13.914	B	Short	Long
–	+	14.821	14.757	14.843	14.878			
+	+	14.88	14.921	14.415	14.932			

- 3.9 A nickel–titanium alloy is used to make components for jet turbine aircraft engines. Cracking is a potentially serious problem in the final part, because it can lead to nonrecoverable failure. A test is run at the parts producer to determine the effect of four factors on cracks. The four factors are pouring temperature ( $A$ ), titanium content ( $B$ ), heat treatment method ( $C$ ), and amount of grain refiner used ( $D$ ). A single replicate of a  $2^4$  design is run, and the lengths (in millimeters) of three cracks induced in a sample coupon subjected to a standard test are measured. The data are shown in Table E3.6.

- (a) Use the average crack length as the response. Estimate the factor effects. Which effects appear to be large?
- (b) Using the results of part (a), form a tentative model and conduct an analysis of variance.
- (c) Analyze the residuals from this model.
- (d) Use the logarithm of the variance of crack length as the response. Is there any evidence that any of the design factors affect the variability in crack length?
- (e) Make recommendations for process operating conditions.

**TABLE E3.6 The Crack Length Experiment from Exercise 3.9**

A	B	C	D	Treatment Combination	Crack Length		
–	–	–	–	(1)	0.011	0.015	0.013
+	–	–	–	<i>a</i>	1.015	1.002	1.018
–	+	–	–	<i>b</i>	0.014	0.012	0.010
+	+	–	–	<i>ab</i>	1.032	1.020	1.046
–	–	+	–	<i>c</i>	0.440	0.444	0.459
+	–	+	–	<i>ac</i>	2.388	2.253	2.467
–	+	+	–	<i>bc</i>	1.228	1.100	1.306
+	+	+	–	<i>abc</i>	2.759	2.560	2.839
–	–	–	+	<i>d</i>	0.010	0.013	0.008
+	–	–	+	<i>ad</i>	0.593	0.542	0.631
–	+	–	+	<i>bd</i>	0.675	0.600	0.752
+	+	–	+	<i>abd</i>	1.362	1.215	1.413
–	–	+	+	<i>cd</i>	0.544	0.591	0.486
+	–	+	+	<i>acd</i>	2.112	2.003	2.218
–	+	+	+	<i>bcd</i>	1.553	1.318	1.625
+	+	+	+	<i>abcd</i>	2.497	2.220	2.753

- 3.10** Reconsider the crack length experiment in Exercise 3.9. Factor *C*, a type of heat treatment, is a categorical variable. Write down the appropriate models for the mean case length response for each type of heat treatment.
- (a) Construct the contour plots of average crack length for these two models. Comment on any differences in the plots.
- (b) Is there any evidence that one heat treatment method results in lower variability in crack length than the other?
- 3.11** An article in *Solid State Technology* (“Orthogonal Design for Process Optimization and its Application in Plasma Etching,” May 1987, pp. 127–132) describes the application of factorial designs in developing a nitride etch process on a single-wafer plasma etcher. The process uses  $C_2F_6$  as the reactant gas. Four factors are of interest: anode–cathode gap (*A*), pressure in the reactor chamber (*B*),  $C_2F_6$  gas flow (*C*), and power applied to the cathode (*D*). The response variable of interest is the etch rate for silicon nitride. A single replicate of a  $2^4$  design is run, and the resulting data are in Table E3.7.
- (a) Estimate the factor effects. Construct a normal probability plot of the factor effects. Which effects appear large?
- (b) Conduct an analysis of variance to confirm your findings for part (a).

**TABLE E3.7 The Etch Rate Experiment from Exercise 3.11**

Run Number	Actual Run Order	A	B	C	D	Etch Rate (Å/min)	Factor	Level	
								Low (-)	High (+)
1	13	-	-	-	-	500	A (cm)	0.8	1.20
2	8	+	-	-	-	669	B (mTorr)	4.5	550
3	12	-	+	-	-	604	C (sccm)	125	200
4	9	+	+	-	-	650	D (W)	275	325
5	4	-	-	+	-	633			
6	15	+	-	+	-	642			
7	16	-	+	+	-	601			
8	3	+	+	+	-	635			
9	1	-	-	-	+	1037			
10	14	+	-	-	+	749			
11	5	-	+	-	+	1052			
12	10	+	+	-	+	868			
13	11	-	-	+	+	1075			
14	2	+	-	+	+	860			
15	7	-	+	+	+	1063			
16	6	+	+	+	+	729			

- (c) Analyze the residuals from this experiment. Comment on the model's adequacy.
- (d) If not all the factors are important, project the  $2^4$  design into a  $2^k$  design with  $k < 4$  and conduct the analysis of variance.
- (e) Draw graphs to interpret any significant interactions.
- (f) Plot the residuals versus the actual run order. What problems might be revealed by this plot?
- 3.12** Consider the single replicate of the  $2^4$  design in Example 3.11. Suppose we had arbitrarily decided to analyze the data assuming that all three- and four-factor interactions were negligible. Conduct this analysis and compare your results with those obtained in the example. Do you think that it is a good idea to arbitrarily assume interactions to be negligible even if they are relatively high-order ones?
- 3.13** An experiment was run in a semiconductor fabrication plant in an effort to increase the number of good chips produced per wafer. Five factors, each at two levels, were studied. The factors (and levels) were  $A$  = aperture setting (small, large),  $B$  = exposure time (20% below nominal),  $C$  = development time (30 sec, 45 sec),  $D$  = mask dimension (small, large), and  $E$  = etch time (14.5 min, 15.5 min). The unreplicated  $2^5$  design shown in Table E3.8 is run.
- (a) Construct a normal probability plot of the effect estimates. Which effects appear to be large?
- (b) Conduct an analysis of variance to confirm your findings for part (a).
- (c) Plot the residuals on normal probability paper. Is the plot satisfactory?
- (d) Plot the residuals versus the predicted yields and versus each of the five factors. Comment on the plots.
- (e) Interpret any significant interactions.



- (f) What are your recommendations regarding process operating conditions?
- (g) Project the  $2^5$  design in this problem into a  $2^k$  design in the important factors. Sketch the design, and show the average and range of yields at each run. Does this sketch aid in data interpretation?
- (h) Construct a contour plot showing how yield changes in terms of the important factors.

**TABLE E3.8 The  $2^5$  Experiment from Exercise 3.13**

(1)	15	<i>d</i>	16	<i>e</i>	14	<i>de</i>	12
<i>a</i>	20	<i>ad</i>	21	<i>ae</i>	23	<i>ade</i>	19
<i>b</i>	70	<i>bd</i>	65	<i>be</i>	71	<i>bde</i>	60
<i>ab</i>	112	<i>abd</i>	100	<i>abe</i>	110	<i>abde</i>	106
<i>c</i>	30	<i>cd</i>	35	<i>ce</i>	31	<i>cde</i>	32
<i>ac</i>	40	<i>acd</i>	45	<i>ace</i>	43	<i>acde</i>	39
<i>bc</i>	79	<i>bcd</i>	85	<i>bce</i>	90	<i>bcde</i>	82
<i>abc</i>	125	<i>abcd</i>	120	<i>abce</i>	130	<i>abcde</i>	128

- 3.14** After running a  $2^4$  factorial design, an engineer obtained the following factor effect estimates:

$$\begin{array}{lll}
 A = -1.3 & AB = -2.9 & ABC = 0.5 \\
 B = 10.1 & AC = 1.9 & ABD = 1.8 \\
 C = 7.5 & AD = -4.0 & ACD = -2.7 \\
 D = -1.4 & BC = -5.1 & BCD = 2.8 \\
 & BD = 2.2 & ABCD = -0.6 \\
 & CD = -0.2 &
 \end{array}$$

Use a normal probability plot to interpret these effects.

- 3.15** Consider the data from the experiment in Exercise 3.11. Suppose that the last run made in this experiment (run order number 16) was lost; in fact, the wafer was broken so that no reading on etch rate could be obtained. One logical approach to analyzing these data is to estimate the missing value that makes the highest-order interaction effect estimate zero. Apply this technique to these data, and compare your results with those obtained in the analysis of the full data set in Exercise 3.10.
- 3.16 Continuation of Exercise 3.15.** Reconsider the situation described in Exercise 3.15.
- (a) Analyze the data by fitting a regression model in the main effects to the 15 observations.
- (b) Investigate the adequacy of this model. What are your conclusions?
- (c) Build a model for the 15 observations with stepwise regression methods, using the main effects and two-factor interactions as the candidate variables.
- (d) Comment on the model obtained above, and compare it to those from Exercise 3.11 and 3.15.
- 3.17 Contour Plots for Exercise 3.11.** Reconsider the etch rate experiment in Exercise 3.11. Construct contour plots for the etch rate response. Suppose that it is important

to operate this process with a mean etch rate very close to 800 Å/m. What operating conditions do you recommend?

- 3.18** An experiment was conducted on a chemical process that produces a polymer. The four factors studied were temperature (*A*), catalyst concentration (*B*), time (*C*), and pressure (*D*). Two responses, molecular weight and viscosity, were observed. The design matrix and the resulting data are given in Table E3.9.
- Consider only the molecular weight response. Plot the effect estimates on a normal probability scale. What effects appear important?
  - Use an analysis of variance to confirm the results from part (a). Is there indication of curvature?
  - Write down a regression model to predict molecular weight as a function of the important variables.
  - Analyze the residuals and comment on model adequacy.
  - Repeat parts (a)–(d) using the viscosity response.

**TABLE E3.9 The Chemical Process Experiment, Exercise 3.18**

Run Number	Actual Run Order					Molecular Weight	Viscosity	Factor	Level	
		<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>				Low (–)	High (+)
1	18	–	–	–	–	2400	1400	<i>A</i> (°C)	100	120
2	9	+	–	–	–	2410	1500	<i>B</i> (%)	4	8
3	13	–	+	–	–	2315	1520	<i>C</i> (min)	20	30
4	8	+	+	–	–	2510	1630	<i>D</i> (psi)	60	75
5	3	–	–	+	–	2615	1380			
6	11	+	–	+	–	2625	1525			
7	14	–	+	+	–	2400	1500			
8	17	+	+	+	–	2750	1620			
9	6	–	–	–	+	2400	1400			
10	7	+	–	–	+	2390	1525			
11	2	–	+	–	+	2300	1500			
12	10	+	+	–	+	2520	1500			
13	4	–	–	+	+	2625	1420			
14	19	+	–	+	+	2630	1490			
15	15	–	+	+	+	2500	1500			
16	20	+	+	+	+	2710	1600			
17	1	0	0	0	0	2515	1500			
18	5	0	0	0	0	2500	1460			
19	16	0	0	0	0	2400	1525			
20	12	0	0	0	0	2475	1500			

- 3.19 Continuation of Exercise 3.18.** Use the regression models for molecular weight and viscosity to answer the following questions.

- Construct a response surface contour plot for molecular weight. In what direction would you adjust the process variables to increase molecular weight?
- Construct a response surface contour plot for viscosity. In what direction would you adjust the process variables to decrease viscosity?

- (c) What operating conditions would you recommend if it is necessary to produce a product with molecular weight between 2400 and 2500 and with low viscosity?
- 3.20** Consider the experiment described in Exercise 3.1. Suppose that only one replicate (four runs) could be obtained in a single 4-hr time period, and the experimenters were concerned about unknown factors that could vary from one time period to another.
- (a) Set up a design in four blocks that will minimize the time effects.
- (b) Analyze the data assuming that the design had been run as in part (a). Compare the results of your analysis with the original analysis in Exercise 3.1.
- (c) Suppose that the experimenter's concerns were realized and that as a result of the time effect, the observations in replicate IV were  $(1) = 34.4$ ,  $a = 44.5$ ,  $b = 34.2$ , and  $ab = 59.9$  instead of the values shown in Exercise 3.1. Reanalyze these new data, assuming that the blocked design from part (a) had been used. How has the time effect influenced your conclusions?
- 3.21** Consider the experiment described in Exercise 3.2. Suppose that this experiment had been run with each replicate considered as a block. Analyze the data and draw conclusions. How do the results of your analysis compare with the analysis of the original Exercise 3.2?
- 3.22** Consider only the data from the first replicate of Exercise 3.2. Set up a design to run these observations in two blocks of four runs each with  $ABC$  confounded with blocks. Analyze the data.
- 3.23** Consider the data from both replicates of Exercise 3.2. Suppose that only four runs could be made under the same conditions, so that it was necessary to run this design in blocks of four runs.
- (a) Set up a design that confounds  $ABC$  in both replicates. Analyze the data.
- (b) Set up a design that confounds  $ABC$  in the first replicate and  $AB$  in the second replicate. Analyze this design, using the data from replicate I to draw conclusions about the  $AB$  effect, and the data from replicate II to draw conclusions about the  $ABC$  effect (this is a design technique called *partial confounding*).
- 3.24** Consider the data from Exercise 3.6. Construct a design with two blocks of eight runs each, with  $ABCD$  confounded with blocks. Analyze the data as if it came from this design.
- 3.25** **Continuation of Exercise 3.24.** Reconsider the situation in Exercise 3.24, and assume that the block effect causes all observations in the second block to be reduced by 20. Analyze the data that result. Is the block effect significant? How are the other conclusions affected?
- 3.26** Consider the situation described in Exercise 3.13. Set up a design in two blocks for 16 observations each of this problem. Analyze the data that result.
- 3.27** Consider the situation described in Exercise 3.13. Set up a design in four blocks of eight runs each, with  $ACDE$  and  $BCE$  confounded with blocks. What is the other confounded effect? Analyze the data and draw conclusions.
- 3.28** Construct a  $2^6$  design in eight blocks of eight runs each. Show the complete set of effects that are confounded with blocks.

- 3.29 Suppose that an experimenter has run a  $2^3$  factorial design (eight trials) in random order. He fits the following model to the data:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \varepsilon$$

- (a) Suppose that the true model is

$$E(y) = \beta_0 + \beta_1x_2 + \beta_2x_2 + \beta_3x_3 + \beta_{12}x_1x_2$$

Does the interaction term result in biased estimates of  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  in the fitted first-order model?

- (b) Suppose that the true model is

$$E(y) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{12}x_1x_2 + \beta_{11}x_1^2$$

Determine the bias induced in the least squares estimates of  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  by the failure to include the terms  $\beta_{12}x_1x_2$  and  $\beta_{11}x_1^2$  in the fitted model.

- 3.30 An experimenter has run a  $2^3$  design in standard order (no randomization). Unknown to the experimenter, a fourth variable  $x_4$  takes on the values shown in Table E3.10.

Suppose that the experimenter fits the first-order model

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \varepsilon$$

Find the bias in the least squares estimates of  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  that results from not including the term  $\beta_4x_4$  in the model.

**TABLE E3.10 The  $2^3$  Experiment for Exercise 3.30**

Run	$x_1$	$x_2$	$x_3$	$x_4$
1	-1	-1	-1	-4
2	1	-1	-1	-3
3	-1	1	-1	-2
4	1	1	-1	-1
5	-1	-1	1	1
6	1	-1	1	2
7	-1	1	1	3
8	1	1	1	4

- 3.31 Consider the runs from a  $2^3$  factorial design shown in Table E3.11, in random order. The first four observations are subject to a block effect of 20 units.

- (a) Find the effect of the block effect on the estimates of the main effects of  $A$ ,  $B$ , and  $C$ .
- (b) Suppose that the design had been set up with  $ABC$  confounded with blocks, and once again, the first 4 runs were subject to a block effect of 20 units. What is the effect of this block effect on the estimates of  $A$ ,  $B$ , and  $C$

**TABLE E3.11** The  $2^3$  Experiment for Exercise 3.31

Run Order	A	B	C	Observation
1	+	+	−	$ab + 20$
2	+	−	−	$a + 20$
3	−	−	−	$(1) + 20$
4	+	+	+	$abc + 20$
5	−	+	−	$b$
6	+	−	+	$ac$
7	−	−	+	$c$
8	−	+	+	$bc$

- 3.32** Consider a  $2^3$  design with a center point that has been replicated  $n_C$  times. Show that the runs at the center do not affect the estimates of any of the factorial affects. Do the center points affect the estimate of the overall mean? Will these results generalize to any  $2^k$  design?
- 3.33** Consider the  $2^k$  design with  $n_C$  center points. Show that an unbiased estimate of the sum of the pure quadratic coefficients ( $\sum_{i=1}^k \beta_{ij}$ ) in a second-order model can be obtained from the difference between  $\bar{y}_F$  and  $\bar{y}_C$ .
- 3.34** Consider the  $2^k$  design with  $n_C$  center points. We gave an  $F$ -test for pure quadratic curvature. Show that the same hypothesis can be tested using a  $t$ -test.
- 3.35** Consider the plasma etching experiment described in Exercise 3.11. Suppose that factor A, the anode–cathode gap, is a hard-to-change factor. Set up an appropriate split-plot design for conducting this experiment.
- 3.36** Consider the chemical process experiment described in Exercise 3.18. Suppose that temperature is a hard-to-change factor. Set up an appropriate split-plot design for conducting this experiment.
- 3.37** The book by Davies (*Design and Analysis of Industrial Experiments*) describes an experiment to study the yield of isatin. The factors studied and their levels are as follows:

Factor	Low (−)	High (+)
A: Acid strength (%)	87	93
B: Reaction time (min)	15	30
C: Amount of acid (mL)	35	45
D: Reaction temperature (°C)	60	70

The data from the 24 factorial is shown in Table E3.12.

- (a) Fit a main-effects-only model to the data from this experiment. Are any of the main effects significant?
- (b) Analyze the residuals. Are there any indications of model inadequacy or violation of the assumptions?

- (c) Find an equation for predicting the yield of isatin over the design space. Express the equation in both coded and engineering units.
- (d) Construct a normal probability plot of the effects and draw conclusions about a tentative model for the data.
- (e) Is there any indication that adding interactions to the model would improve the results that you have obtained?

**TABLE E3.12** The  $2^4$  factorial Experiment for Exercise 3.27

A	B	C	D	Y (Yield)
-1	-1	-1	-1	6.08
1	-1	-1	-1	6.04
-1	1	-1	-1	6.53
1	1	-1	-1	6.43
-1	-1	1	-1	6.31
1	-1	1	-1	6.09
-1	1	1	-1	6.12
1	1	1	-1	6.36
-1	-1	-1	1	6.79
1	-1	-1	1	6.68
-1	1	-1	1	6.73
1	1	-1	1	6.08
-1	-1	1	1	6.77
1	-1	1	1	6.38
-1	1	1	1	6.49
1	1	1	1	6.23

- 3.38** An article in *Quality and Reliability Engineering International* (2010, Vol. 26, pp. 223–233) presents a  $2^5$  factorial design. The experiment is shown in Table E3.13.
- (a) Analyze the data from this experiment. Identify the significant factors and interactions.
  - (b) Analyze the residuals from this experiment. Are there any indications of model inadequacy or violations of the assumptions?
  - (c) One of the factors from this experiment does not seem to be important. If you drop this factor, what type of design remains? Analyze the data using the full factorial model for only the four active factors. Compare your results with those obtained in part (a).
  - (d) Find settings of the active factors that maximize the predicted response.
- 3.39** Suppose that the  $2^5$  experiment in Exercise 3.38 had been run in two blocks with ABCDE confounded with blocks.
- (a) Set up the design.

TABLE E3.13 The  $2^5$  Experiment for Exercise 3.38

A	B	C	D	E	y
-1	-1	-1	-1	-1	8.11
1	-1	-1	-1	-1	5.56
-1	1	-1	-1	-1	5.77
1	1	-1	-1	-1	5.82
-1	-1	1	-1	-1	9.17
1	-1	1	-1	-1	7.8
-1	1	1	-1	-1	3.23
1	1	1	-1	-1	5.69
-1	-1	-1	1	-1	8.82
1	-1	-1	1	-1	14.23
-1	1	-1	1	-1	9.2
1	1	-1	1	-1	8.94
-1	-1	1	1	-1	8.68
1	-1	1	1	-1	11.49
-1	1	1	1	-1	6.25
1	1	1	1	-1	9.12
-1	-1	-1	-1	1	7.93
1	-1	-1	-1	1	5
-1	1	-1	-1	1	7.47
1	1	-1	-1	1	12
-1	-1	1	-1	1	9.86
1	-1	1	-1	1	3.65
-1	1	1	-1	1	6.4
1	1	1	-1	1	11.61
-1	-1	-1	1	1	12.43
1	-1	-1	1	1	17.55
-1	1	-1	1	1	8.87
1	1	-1	1	1	25.38
-1	-1	1	1	1	13.06
1	-1	1	1	1	18.85
-1	1	1	1	1	11.78
1	1	1	1	1	26.05

(b) Analyze the blocked design. How large is the block effect? Compare your conclusions with the completely randomized design.

**3.40** Rework Exercise 3.39, but now suppose that all of the observations in one of the two blocks are increased by 5 units. This introduces a fairly large block effect. Compare your results with those obtained from the completely randomized design.

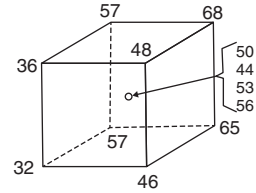
**3.41** Suppose that the  $2^5$  experiment in Exercise 3.38 had been run in four blocks.

(a) Set up the design. Carefully show which effects are confounded with blocks.

(b) Analyze the blocked design. How large is the block effect? Compare your conclusions with the completely randomized design.

3.42 Consider the 2<sup>3</sup> experiment shown in the figure below:

Process Variables				Coded Variables			Yield, y
Run	Temp (°C)	Pressure (psing)	Conc (g/l)	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	
1	120	40	15	-1	-1	-1	32
2	160	80	15	1	-1	-1	46
3	120	40	15	-1	1	-1	57
4	150	80	15	1	1	-1	65
5	120	40	30	-1	-1	1	36
6	150	80	30	1	-1	1	48
7	120	40	30	-1	1	1	57
8	150	80	30	1	1	1	68
9	140	60	22.5	0	0	0	50
10	140	60	22.5	0	0	0	44
11	140	60	22.5	0	0	0	53
12	140	60	22.5	0	0	0	56

$$x_1 = \frac{\text{Temp} - 140}{20}, x_2 = \frac{\text{Pressure} - 60}{20}, x_3 = \frac{\text{Conc} - 22.5}{7.5}$$


When running a designed experiment, it is sometimes difficult to reach and hold the precise factor levels required by the design. Small discrepancies are not important, but large ones are potentially of more concern. To illustrate, the experiment presented in Table E3.14 shows a variation of the 2<sup>3</sup> design above, where many of the test combinations are not exactly the ones specified in the design. Most of the difficulty seems to have occurred with the temperature variable.

Fit a first-order model to both the original data and the data in Table E3.14. Compare the inference from the two models. What conclusions can you draw from this simple example?

TABLE E3.14 Revised Experimental Data

Process Variables				Coded Variables			Yield, y
Run	Temp (°C)	Pressure (psing)	Conc (g/l)	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	
1	125	41	14	-0.75	-0.95	-1.133	32
2	158	40	15	0.90	-1	-1	46
3	121	82	15	-0.95	1.1	-1	57
4	160	80	15	1	1	-1	65
5	118	39	33	-1.10	-1.05	1.14	36
6	163	40	30	1.15	-1	1	48
7	122	80	30	-0.90	1	1	57
8	165	83	30	1.25	1.15	1	68
9	140	60	22.5	0	0	0	50
10	140	60	22.5	0	0	0	44
11	140	60	22.5	0	0	0	53
12	140	60	22.5	0	0	0	56



- 3.43** Suppose that there are three factor each at two levels. The experimenter plans to fit a main effects plus two-factor interaction model and can use exactly 12 runs. Find a D-optimal design for this problem.
- 3.44** Reconsider the situation in Exercise 3.43. Suppose that the experimenter decided to run a single replicate of the  $2^3$  design plus four center points.
- (a) Is this a D-optimal design?
- (b) Compare the prediction performance of this design to the D-optimal design from Exercise 3.43 at the corners of the cube.
- 3.45** Consider the  $2^5$  design in Exercise 3.38. Suppose that run *abe* was missing. Analyze the remaining 15 runs and compare your results with those obtained in Exercise 3.38.

### Short Answer Questions

- 3.46** The key concept of the phrase “Block if you can, randomize if you can’t.” means that:
- (a) It is usually better to not randomize within blocks.
- (b) Blocking violates the assumption of constant variance.
- (c) Create blocks by using each level of the nuisance factor as a block and randomize within blocks.
- (d) Randomizing the runs is preferable to randomizing blocks.
- 3.47** Suppose that a  $2^3$  factorial experiment with 2 replicates has been conducted. The experiment has been conducted in blocks. The error sum of squares is 500 and the block sum of squares is 250. If the experiment had been conducted as a completely randomized design the estimate of the error variance  $\sigma^2$  would be
- (a) 25.0
- (b) 25.5
- (c) 35.0
- (d) 38.5
- (e) None of the above
- 3.48** Physics graduate student Laura Van Ertia has conducted a  $2^2$  factorial experiment as a CRD (completely randomized design), hoping to solve the mystery of the unified theory and complete her dissertation (good luck on that). The results of this experiment are summarized in the following ANOVA table:

Source	SS	DF	MS	F
Model	X	X	28.57	
A	X	1	25.00	
B	75.00	X	75.00	
AB	50.00	1	X	
Error	X	X	X	
Total	250.00	15		

Answer the following questions about this experiment.

- (a) Fill in the blanks in the ANOVA table.
- (b) The number of degrees of freedom for error is \_\_\_\_\_
- (c) The estimate of  $\sigma^2$  is \_\_\_\_\_
- (d) The value of the test statistic for factor A is \_\_\_\_\_
- (e) If the significance level is 0.01, your conclusions are to reject the null hypothesis that factor A is not significant (answer either Yes or No) \_\_\_\_\_

- 3.49 Suppose that the experimenters from Exercise 3.37 fit the following model in the four main effects (model shown in coded units):

$$\hat{y} = 0.25 + 1.5x_1 - 2.2x_2 + 3.0x_3 - 1.1x_4$$

Find the predicted value of the response at the following points:

- (a) A = 89, B = 20, C = 38, D = 66
  - (b) A = 90, B = 16, C = 40, D = 70
  - (c) A = 87, B = 28, D = 42, D = 61
  - (d) A = 90, B = 27, D = 37, D = 69
- 3.50 You are conducting a  $2^6$  factorial design in 4 blocks. Which of the runs below are in the same block as *abc*?
- (a) *a*
  - (b) *abcd*
  - (c) *cdf*
  - (d) *bcdef*
  - (e) All of the above
  - (f) None of the above
- 3.51 In a randomized complete block design the number of treatments and blocks must be equal. **True False**
- 3.52 The square root of the residual mean square in the ANOVA estimates the standard deviation of an observation from any factor level. **True False**
- 3.53 In a randomized complete block design, the blocks are always chosen randomly. **True False**
- 3.54 There is no effective way to check the validity of the assumption of normality of the observations in a designed experiment. **True False**
- 3.55 An interaction effect in the model from a factorial experiment is a way of incorporating curvature into the response surface model representation of the results. **True False**

---

# 4

---

## TWO-LEVEL FRACTIONAL FACTORIAL DESIGNS

### 4.1 INTRODUCTION

As the number of factors in a  $2^k$  factorial design increases, the number of runs required for a complete replicate of the design rapidly outgrows the resources of most experimenters. For example, a complete replicate of the  $2^6$  design requires 64 runs. In this design, only 6 of the 63 degrees of freedom are used to estimate the main effects, and only 15 degrees of freedom are used to estimate the two-factor interactions. The remaining 42 degrees of freedom are associated with three-factor and higher interactions.

If the experimenter can reasonably assume that certain high-order interactions are negligible, then information on the main effects and low-order interactions may be obtained by running only a fraction of the complete factorial experiment. These **fractional factorial designs** are among the most widely used types of design in industry.

A major use of fractional factorials is in **screening experiments**. These are experiments in which many factors are considered with the purpose of identifying those factors (if any) that have large effects. Remember that screening experiments are usually performed early in a response surface study when it is likely that many of the factors initially considered have little or no effect on the response. The factors that are identified as important are then investigated more thoroughly in subsequent experiments.

The successful use of fractional factorial designs is based on three key ideas:

1. *The Sparsity-of-Effects Principle*. When there are several variables, the system or process is likely to be driven primarily by some of the main effects and low-order interactions.

**TABLE 4.1 Plus and Minus Signs for the  $2^3$  Factorial Design**

Treatment Combination	Factorial Effect							
	<i>I</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>AB</i>	<i>AC</i>	<i>BC</i>	<i>ABC</i>
<i>a</i>	+	+	-	-	-	-	+	+
<i>b</i>	+	-	+	-	-	+	-	+
<i>c</i>	+	-	-	+	+	-	-	+
<i>abc</i>	+	+	+	+	+	+	+	+
<i>ab</i>	+	+	+	-	+	-	-	-
<i>ac</i>	+	+	-	+	-	+	-	-
<i>bc</i>	+	-	+	+	-	-	+	-
(1)	+	-	-	-	+	+	+	-

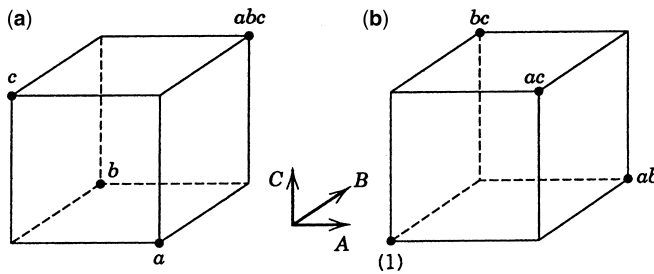
2. *The Projection Property.* Fractional factorial designs can be projected into stronger (larger) designs in the subset of significant factors.
3. *Sequential Experimentation.* It is possible to combine the runs of two (or more) fractional factorials to assemble sequentially a larger design to estimate the factor effects and interactions of interest.

We will focus on these principles in this chapter and illustrate them with several examples.

### 4.2 THE ONE-HALF FRACTION OF THE $2^k$ DESIGN

Consider the situation in which three factors, each at two levels, are of interest, but the experimenters do not wish to run all  $2^3 = 8$  treatment combinations. Suppose they consider a design with four runs. This suggests a **one-half fraction** of a  $2^3$  design. Because the design contains  $2^{3-1} = 4$  treatment combinations, a one-half fraction of the  $2^3$  design is often called a  $2^{3-1}$  design.

The table of plus and minus signs for the  $2^3$  design is shown in Table 4.1. Suppose we select the four treatment combinations *a*, *b*, *c*, and *abc* as our one-half fraction. These runs are shown in the top half of Table 4.1 and in Fig. 4.1a. Notice that the  $2^{3-1}$  design is formed by selecting only those treatment combinations that have a plus in the *ABC* column. Thus, *ABC* is called the **generator** of this particular fraction. Sometimes we will refer to a



**Figure 4.1** The two one-half fractions of the  $2^3$  design. (a) The principal fraction,  $I = +ABC$ . (b) The alternate fraction,  $I = -ABC$ .

generator such as  $ABC$  as a **word**. Furthermore, the identity column  $I$  is also always plus, so we call

$$I = ABC$$

the **defining relation** for our design. In general, the defining relation for a fractional factorial will be the set of all columns that are equal to the identity column  $I$ .

The treatment combinations in the  $2^{3-1}$  design yield three degrees of freedom that we may use to estimate the main effects. Referring to Table 4.1, we note that the linear combinations of the observations used to estimate the main effects of  $A$ ,  $B$ , and  $C$  are

$$[A] = \frac{1}{2}(a - b - c + abc)$$

$$[B] = \frac{1}{2}(-a + b - c + abc)$$

$$[C] = \frac{1}{2}(-a - b + c + abc)$$

It is also easy to verify that the linear combinations of the observations used to estimate the two-factor interactions are

$$[BC] = \frac{1}{2}(a - b - c + abc)$$

$$[AC] = \frac{1}{2}(-a + b - c + abc)$$

$$[AB] = \frac{1}{2}(-a - b + c + abc)$$

Thus,  $[A] = [BC]$ ,  $[B] = [AC]$ , and  $[C] = [AB]$ ; consequently, it is impossible to differentiate between  $A$  and  $BC$ ,  $B$  and  $AC$ , and  $C$  and  $AB$ . In fact, when we estimate  $A$ ,  $B$ , and  $C$  we are really estimating  $A + BC$ ,  $B + AC$ , and  $C + AB$ . Two or more effects that have this property are called **aliases**. In our example,  $A$  and  $BC$  are aliases,  $B$  and  $AC$  are aliases, and  $C$  and  $AB$  are aliases. We indicate this by the notation  $[A] \rightarrow A + BC$ ,  $[B] \rightarrow B + AC$ , and  $[C] \rightarrow C + AB$ .

The alias structure for this design may be easily determined by using the defining relation  $I = ABC$ . Multiplying any column by the defining relation yields the aliases for that effect. In our example, this yields as the alias of  $A$

$$A \cdot I = A \cdot ABC = A^2BC$$

or, because the square of any column is just the identity  $I$ ,

$$A = BC$$

Similarly, we find the aliases of  $B$  and  $C$  as

$$B \cdot I = B \cdot ABC$$

$$B = AB^2C = AC$$

and

$$C \cdot I = C \cdot ABC$$

$$C = ABC^2 = AB$$

This one-half fraction, with  $I = +ABC$ , is usually called the **principal fraction**.

Now suppose that we had selected the other one-half fraction; that is, the treatment combinations in Table 4.1 associated with a minus sign in the  $ABC$  column. This **alternate** or **complementary fraction** is shown in Fig. 4.1b. Notice that it consists of the runs labeled (1),  $ab$ ,  $ac$ , and  $bc$ . The defining relation for this design is

$$I = -ABC$$

The linear combination of the observations, say  $[A]'$ ,  $[B]'$ , and  $[C]'$ , from the alternate fraction gives us the following alias relationships:

$$[A]' \longrightarrow A - BC$$

$$[B]' \longrightarrow B - AC$$

$$[C]' \longrightarrow C - AB$$

Thus, when we estimate  $A$ ,  $B$ , and  $C$  with this particular fraction, we are really estimating  $A - BC$ ,  $B - AC$ , and  $C - AB$ .

In practice, it does not matter which fraction is actually used. Both fractions belong to the same **family**; that is, the two one-half fractions form a complete  $2^3$  design. This is easily seen by referring to Fig. 4.1a and 4.1b.

Suppose that after running one of the one-half fractions of the  $2^3$  design, the other one was also run. Thus, all eight runs associated with the full  $2^3$  are now available. We may now obtain de-aliased estimates of all the effects by analyzing the eight runs as a full  $2^3$  design in two blocks of four runs each. This could also be done by adding and subtracting the linear combination of effects from the two individual fractions. For example, consider  $[A] \rightarrow A + BC$  and  $[A]' \rightarrow A - BC$ . This implies that

$$\frac{1}{2}([A] + [A]') = \frac{1}{2}(A + BC + A - BC) \longrightarrow A$$

and

$$\frac{1}{2}([A] - [A]') = \frac{1}{2}(A + BC - A + BC) \longrightarrow BC$$

Thus, for all three pairs of linear combinations, we would obtain the following:

$i$	From $\frac{1}{2}([i] + [i]')$	From $\frac{1}{2}([i] - [i]')$
$A$	$A$	$BC$
$B$	$B$	$AC$
$C$	$C$	$AB$

**Design Resolution** The preceding  $2^{3-1}$  design is called a **resolution III design**. In such a design, main effects are aliased with two-factor interactions. A design is of resolution  $R$  if no  $p$ -factor effect is aliased with another effect containing less than  $R - p$  factors. We usually employ a Roman numeral subscript to denote design resolution; thus, the one-half fraction of the  $2^3$  design with the defining relation  $I = ABC$  (or  $I = -ABC$ ) is a  $2^{3-1}_{III}$  design.

Designs of resolution III, IV, and V are particularly important. The definitions of these designs and an example of each follow.



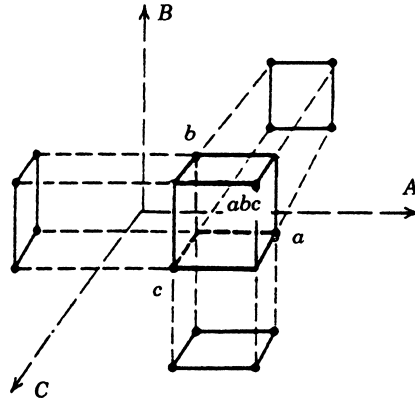


Figure 4.2 Projection of a  $2^{3-1}_{\text{III}}$  design into three  $2^2$  designs.

**Projection of Fractions into Factorials** Any fractional factorial design of resolution  $R$  contains complete factorial designs (possibly replicated factorials) in any subset of  $R - 1$  factors. This is an important and useful concept. For example, if an experimenter has several factors of potential interest but believes that only  $R - 1$  of them have important effects, then a fractional factorial design of resolution  $R$  is the appropriate choice of design. If the experimenter is correct, then the fractional factorial design of resolution  $R$  will project into a full factorial in the  $R - 1$  significant factors. This process is illustrated in Fig. 4.2 for the  $2^{3-1}_{\text{III}}$  design, which projects into a  $2^2$  design in every subset of two factors.

Because the maximum possible resolution of a one-half fraction of the  $2^k$  design is  $R = k$ , every  $2^{k-1}$  design will project into a full factorial in any  $k - 1$  of the original  $k$  factors. Furthermore, a  $2^{k-1}$  design may be projected into two replicates of a full factorial in any subset of  $k - 2$  factors, four replicates of a full factorial in any subset of  $k - 3$  factors, and so on.

**Sequences of Fractional Factorials** Using fractional factorial designs often leads to great economy and efficiency in experimentation, particularly if the runs can be made sequentially. For example, suppose that we were investigating  $k = 4$  factors ( $2^4 = 16$  runs). It is almost always preferable to run a  $2^{4-1}_{\text{IV}}$  fractional design (8 runs), analyze the results, and then decide on the best set of runs to perform next. If it is necessary to resolve ambiguities, we can always run the alternate fraction and complete the  $2^4$  design. When this method is used to complete the design, both one-half fractions represent blocks of the complete design with the highest-order interaction confounded with blocks (here  $ABCD$  would be confounded). Thus, sequential experimentation has the result of losing information only on the highest-order interaction. Alternatively, in many cases we learn enough from the one-half fraction to proceed to the next stage of experimentation, which might involve adding or removing factors, changing responses, or varying some of the factors over new ranges. This potential saving can prove valuable, with additional resources being available for subsequent stages of experimentation.

**Example 4.1 The Filtration Rate Experiment** Consider the filtration rate experiment in Example 3.2. The original design, shown in Table 3.9, is a single replicate of the  $2^4$  design. In this example, we found that the main effects  $A$ ,  $C$ , and  $D$  and the interactions



**TABLE 4.3** The  $2_{IV}^{4-1}$  Design with the Defining Relation  $I = ABCD$

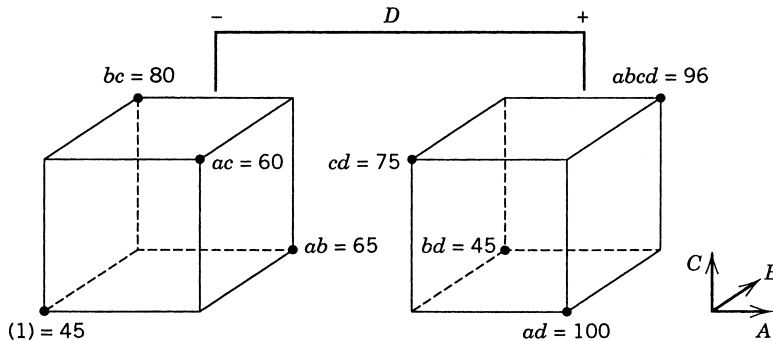
Run	Basic Design			$D = ABC$	Treatment Combination	Filtration Rate
	A	B	C			
1	-	-	-	-	(1)	45
2	+	-	-	+	<i>ad</i>	100
3	-	+	-	+	<i>bd</i>	45
4	+	+	-	-	<i>ab</i>	65
5	-	-	+	+	<i>cd</i>	75
6	+	-	+	-	<i>ac</i>	60
7	-	+	+	-	<i>bc</i>	80
8	+	+	+	+	<i>abcd</i>	96

$AC$  and  $AD$  were different from zero. We will now return to this experiment and simulate what would have happened if a half-fraction of the  $2^4$  design had been run instead of the full factorial.

We will use the  $2^{4-1}$  design with  $I = ABCD$ , because this choice of generator will result in a design of the highest possible resolution (IV). To construct the design, we first write down the basic design, which is a  $2^3$  design, as shown in the first three columns of Table 4.3. Because the generator  $ABCD$  is positive, this  $2_{IV}^{4-1}$  design is the principal fraction. The design is shown graphically in Fig. 4.3.

Using the defining relation, we note that each main effect is aliased with a three-factor interaction; that is,  $A = A^2BCD = BCD$ ,  $B = AB^2CD = ACD$ ,  $C = ABC^2D = ABD$ , and  $D = ABCD^2 = ABC$ . Furthermore, every two-factor interaction is aliased with another two-factor interaction. These alias relationships are  $AB = CD$ ,  $AC = BD$ , and  $BC = AD$ . The four main effects plus the three two-factor interaction alias pairs account for the seven degrees of freedom for the design.

At this point, we would normally randomize the eight runs and perform the experiment. Because we have already run the full  $2^4$  design, we will simply select the eight observed filtration rates from Example 3.2 that correspond to the runs in the  $2_{IV}^{4-1}$  design. These observations are shown in the last column of Table 4.3 and are also shown in Fig. 4.3.



**Figure 4.3** The  $2_{IV}^{4-1}$  design for the filtration rate experiment of Example 4.1.

**TABLE 4.4 Estimates of Effects and Aliases from Example 4.1**

Estimate	Alias Structure <sup>a</sup>
[A] = 19.00	[A] → <b>A</b> + BCD
[B] = 1.50	[B] → B + ACD
[C] = 14.00	[C] → <b>C</b> + ABD
[D] = 16.50	[D] → <b>D</b> + ABC
[AB] = -1.00	[AB] → AB + CD
[AC] = -18.50	[AC] → <b>AC</b> + BD
[BC] = 19.00	[BC] → BC + <b>AD</b>

<sup>a</sup>Significant effects are shown in boldface type.

The estimates of the effects obtained from this  $2^{4-1}_{IV}$  design are shown in Table 4.4. To illustrate the calculations, the linear combination of observations associated with the A effect is

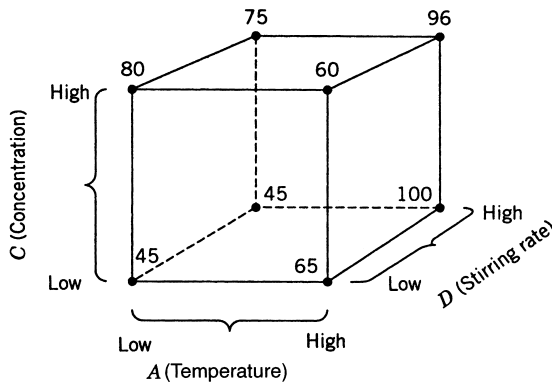
$$[A] = \frac{1}{4}(-45 + 100 - 45 + 65 - 75 + 60 - 80 + 96) = 19.00 \rightarrow A + BCD$$

whereas for the AB effect we obtain

$$[AB] = \frac{1}{4}(45 - 100 - 45 + 65 + 75 - 60 - 80 + 96) = -1.00 \rightarrow AB + CD$$

From inspection of the information in Table 4.4, it is not unreasonable to conclude that the main effects A, C, and D are large and that the AC and AD interactions are also significant. This agrees with the conclusions from the analysis of the complete  $2^4$  design in Example 3.2.

Because factor B is not significant, we may drop it from consideration. Consequently, we may project this  $2^{4-1}_{IV}$  design into a single replicate of the  $2^3$  design in factors A, C, and D, as shown in Fig. 4.4. Visual examination of this cube plot makes us more comfortable with the conclusions reached above. Notice that if the temperature (A) is at the low level, the concentration (C) has a large positive effect, whereas if the temperature is at the high level, the concentration has a very small effect. This is likely due to an AC interaction.



**Figure 4.4** Projection of the  $2^{4-1}_{IV}$  design into  $2^3$  design in A, C, and D for Example 4.1.

Furthermore, if the temperature is at the low level, the effect of the stirring rate ( $D$ ) is negligible, whereas if the temperature is at the high level, the stirring rate has a large positive effect. This is likely due to the  $AD$  interaction tentatively identified previously.

Now suppose that the experimenter decided to run the alternate fraction, given by  $I = -ABCD$ . It is straightforward to show that the design and the responses are as follows:

Run	$A$	$B$	$C$	$D = -ABC$	Treatment Combination	Filtration Rate
1	-	-	-	+	$d$	43
2	+	-	-	-	$a$	71
3	-	+	-	-	$b$	48
4	+	+	-	+	$abd$	104
5	-	-	+	-	$c$	68
6	+	-	+	+	$acd$	86
7	-	+	+	+	$bcd$	70
8	+	+	+	-	$abc$	65

The linear combinations of observations obtained from this alternate fraction are

$$[A]' = 24.25 \longrightarrow A - BCD$$

$$[B]' = 4.75 \longrightarrow B - ACD$$

$$[C]' = 5.75 \longrightarrow C - ABD$$

$$[D]' = 12.75 \longrightarrow D - ABC$$

$$[AB]' = 1.25 \longrightarrow AB - CD$$

$$[AC]' = -17.75 \longrightarrow AC - BD$$

$$[BC]' = -14.25 \longrightarrow BC - AD$$

These estimates may be combined with those obtained from the original one-half fraction to yield the following estimates of the effects:

$i$	From $\frac{1}{2}([A] + [A]')$	From $\frac{1}{2}([A] - [A]')$
$A$	21.63 $\longrightarrow A$	-2.63 $\longrightarrow BCD$
$B$	3.13 $\longrightarrow B$	-1.63 $\longrightarrow ACD$
$C$	9.88 $\longrightarrow C$	4.13 $\longrightarrow ABD$
$D$	14.63 $\longrightarrow D$	1.88 $\longrightarrow ABC$
$AB$	0.13 $\longrightarrow AB$	-1.13 $\longrightarrow CD$
$AC$	-18.13 $\longrightarrow AC$	-0.38 $\longrightarrow BD$
$BC$	2.38 $\longrightarrow BC$	16.63 $\longrightarrow AD$

These estimates agree exactly with those from the original analysis of the data as a single replicate of a  $2^4$  factorial design, as reported in Example 3.2.

TABLE 4.5 A  $2^{5-1}$  Design for Example 4.2

Run	A	B	C	D	$E = ABCD$	Treatment Combination	Resistivity
1	-	-	-	-	+	<i>e</i>	15.1
2	+	-	-	-	-	<i>a</i>	20.6
3	-	+	-	-	-	<i>b</i>	68.7
4	+	+	-	-	+	<i>abe</i>	101.0
5	-	-	+	-	-	<i>c</i>	32.9
6	+	-	+	-	+	<i>ace</i>	46.1
7	-	+	+	-	+	<i>bce</i>	87.5
8	+	+	+	-	-	<i>abc</i>	119.0
9	-	-	-	+	-	<i>d</i>	11.3
10	+	-	-	+	+	<i>ade</i>	19.6
11	-	+	-	+	+	<i>bde</i>	62.1
12	+	+	-	+	-	<i>abd</i>	103.2
13	-	-	+	+	+	<i>cde</i>	27.1
14	+	-	+	+	-	<i>acd</i>	40.3
15	-	+	+	+	-	<i>bcd</i>	87.7
16	+	+	+	+	+	<i>abcde</i>	128.3

**Example 4.2 A  $2^{5-1}$  Design** Five factors in a manufacturing process for an integrated circuit were investigated in a  $2^{5-1}$  design with the objective of learning how these factors affect the resistivity of the wafer. The five factors were  $A$  = implant dose,  $B$  = temperature,  $C$  = time,  $D$  = oxide thickness, and  $E$  = furnace position. Each factor was run at two levels. The construction of the  $2^{5-1}$  design is shown in Table 4.5. Notice that the design was constructed by writing down the basic design having 16 runs (a  $2^4$  design in  $A$ ,  $B$ ,  $C$ , and  $D$ ), selecting  $ABCDE$  as the generator, and then setting the levels of the fifth factor  $E = ABCD$ . Figure 4.5 gives a pictorial representation of the design.

The defining relation for the design is  $I = ABCDE$ . Consequently, every main effect is aliased with a four-factor interaction:

$$[A] \longrightarrow A + BCDE$$

$$[B] \longrightarrow B + ACDE$$

$$[C] \longrightarrow C + ABDE$$

$$[D] \longrightarrow D + ABCE$$

$$[E] \longrightarrow E + ABCD$$

Every two-factor interaction is aliased with a three-factor interaction:

$$[AB] \longrightarrow AB + CDE \quad [BD] \longrightarrow BD + ACE$$

$$[AC] \longrightarrow AC + BDE \quad [BE] \longrightarrow BE + ACD$$

$$[AD] \longrightarrow AD + BCE \quad [CD] \longrightarrow CD + ABE$$

$$[AE] \longrightarrow AE + BCD \quad [CE] \longrightarrow CE + ABD$$

$$[BC] \longrightarrow BC + ADE \quad [DE] \longrightarrow DE + ABC$$

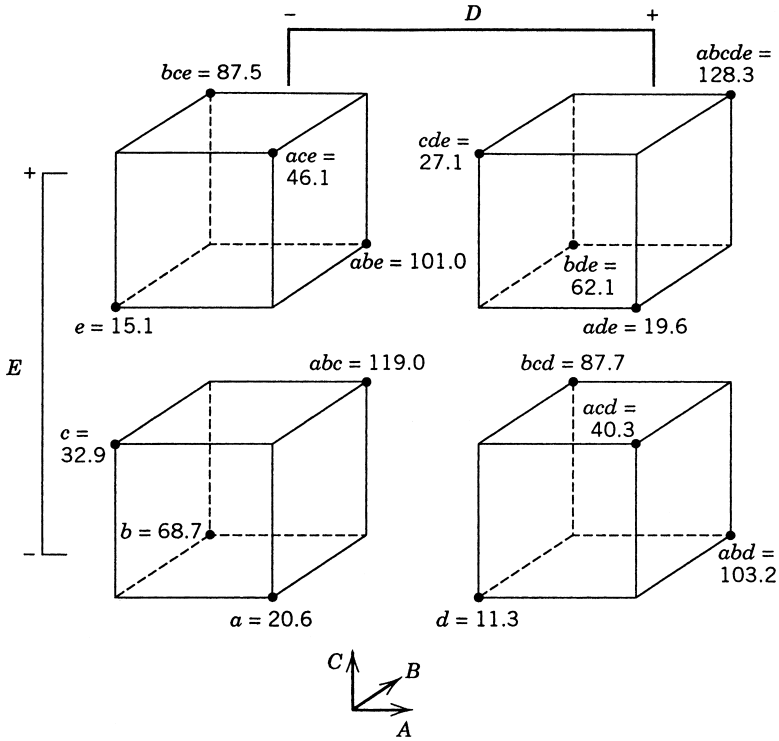


Figure 4.5 The  $2^{5-1}_V$  design for Example 4.2.

This is a resolution V design. We would expect this  $2^{5-1}_V$  design to provide excellent information concerning the main effects and all two-factor interactions.

The estimates of the effects are

$$\begin{aligned}
 [A] &\longrightarrow A + BCDE = 23.2125 \\
 [B] &\longrightarrow B + ACDE = 68.0625 \\
 [C] &\longrightarrow C + ABDE = 20.9125 \\
 [D] &\longrightarrow D + ABCE = -1.4125 \\
 [E] &\longrightarrow E + ABCD = 0.3875 \\
 [AB] &\longrightarrow AB + CDE = 13.1625 & [BD] &\longrightarrow BD + ACE = 2.6875 \\
 [AC] &\longrightarrow AC + BDE = 1.4125 & [BE] &\longrightarrow BE + ACD = -0.3125 \\
 [AD] &\longrightarrow AD + BCE = 2.5875 & [CD] &\longrightarrow CD + ABE = 0.8875 \\
 [AE] &\longrightarrow AE + BCD = 2.5875 & [CE] &\longrightarrow CE + ABD = 1.8875 \\
 [BC] &\longrightarrow BC + ADE = 0.9675 & [DE] &\longrightarrow DE + ABC = -1.7375
 \end{aligned}$$

The effects A, B, C, and AB seem large. Figure 4.6 presents a normal probability plot of the effect estimates from this experiment. This plot confirms that the main effects of A, B,

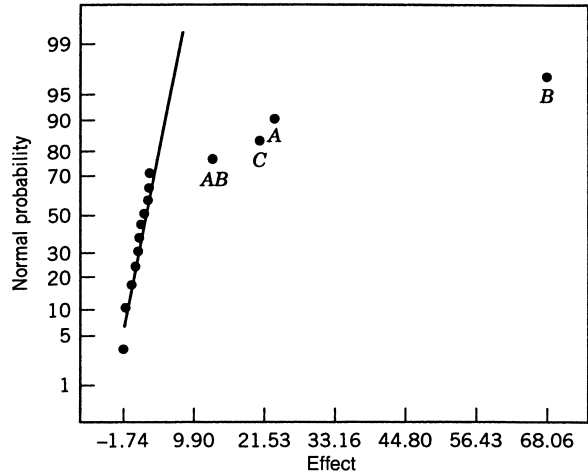


Figure 4.6 Normal probability plot of effects for Example 4.2.

and *C*, and the *AB* interaction are large. Remember that, because of aliasing, these effects are really  $A + BCDE$ ,  $B + ACDE$ ,  $C + ABDE$ , and  $AB + CDE$ . However, because it seems plausible that three-factor and higher interactions are negligible, we feel safe in concluding that only *A*, *B*, *C*, and *AB* are important effects.

Table 4.6 summarizes the analysis of variance for this experiment. The model sum of squares is  $SS_{\text{model}} = SS_A + SS_B + SS_C + SS_{AB} = 23,127.63$ , and this accounts for over 99% of the total variability in the response. The regression model that would be used to predict resistivity over the space of the three factors *A*, *B*, and *C* is

$$\hat{y} = 60.65625 + 11.60625x_1 + 34.03125x_2 + 10.45625x_3 + 6.58125x_1x_2$$

where  $x_1$ ,  $x_2$ , and  $x_3$  are coded variables on the interval  $-1, +1$  that represent *A*, *B*, and *C*.

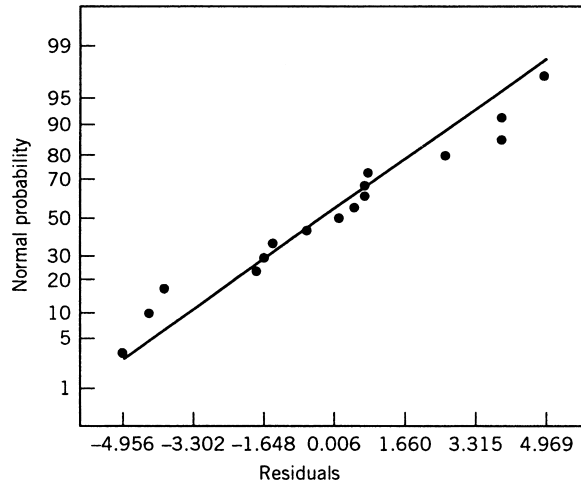
Figure 4.7 presents a normal probability plot of the residuals from this model, and Fig. 4.8 is a plot of the residuals versus the predicted values. Both plots are satisfactory.

The three factors *A*, *B*, and *C* have large positive effects. The *AB* or implant dose–temperature interaction is plotted in Fig. 4.9. This plot indicates that the implant dose has a much stronger effect on resistivity when temperature is at the high level.

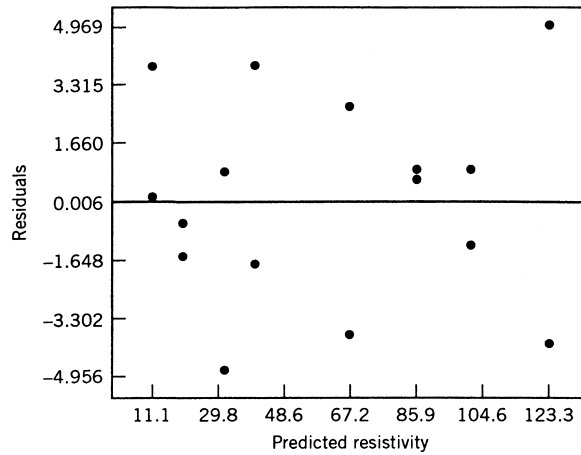
The  $2^{5-1}$  design will collapse into two replicates of a  $2^3$  design in any three of the original five factors. (Looking at Fig. 4.5 will help you visualize this.) Figure 4.10 is a cube

TABLE 4.6 Analysis of Variance for Example 4.2

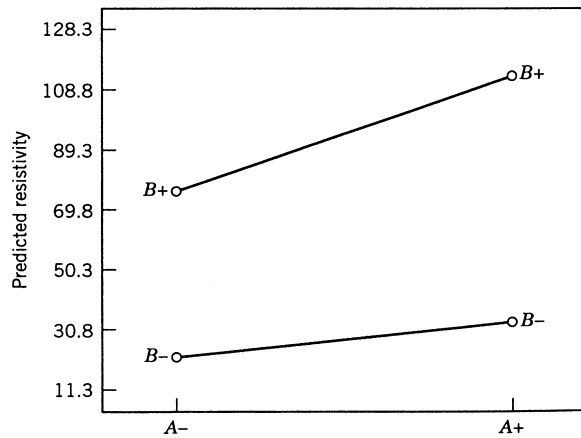
Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	$F_0$
<i>A</i>	2155.28	1	2155.28	193.20
<i>B</i>	18,530.02	1	18,530.02	1791.24
<i>C</i>	1749.33	1	1749.33	184.61
<i>AB</i>	693.01	1	693.01	73.78
Error	132.59	11	12.05	
Total	23,260.22	15		



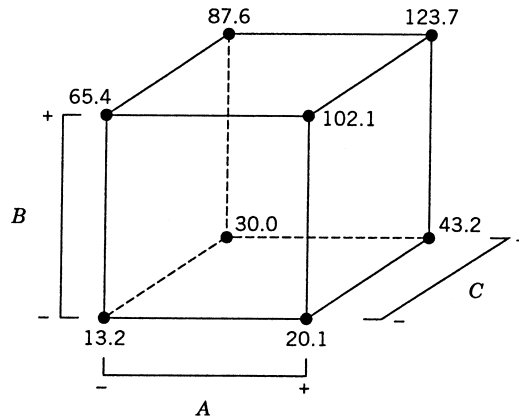
**Figure 4.7** Normal probability plot of the residuals for Example 4.2.



**Figure 4.8** Plot of residuals versus predicted resistivity, Example 4.2.



**Figure 4.9** Implant dose-temperature interaction, Example 4.2.



**Figure 4.10** Projection of the  $2^{5-1}_{\text{V}}$  design in Example 4.2 into two replicates of a  $2^3$  design in the factors  $A$ ,  $B$ , and  $C$ .

plot in the factors  $A$ ,  $B$ , and  $C$  with the average resistivity superimposed on the eight corners. It is clear from inspection of the cube plot that highest resistivity values are achieved with  $A$ ,  $B$ , and  $C$  all at the high level. Factors  $D$  and  $E$  have little effect on resistivity and may be set to values that optimize other objectives (such as cost).

**JMP Analysis of Example 4.2** We can also analyze the fractional factorial design in Example 4.2 using the screening platform in JMP. Table 4.7 contains the output. The upper part of the table shows the results from applying Lenth's method to the data from the experiment. The simulated  $P$ -values indicate that the same factors that we identified in Example 4.1, the main effects of  $A$ ,  $B$ , and  $C$ , and the two-factor  $AB$  interaction are active. The bottom portion of the table contains a half-normal probability plot of the effects. Table 4.8 is the results from fitting this model to the data. This table also includes the final prediction equation, the prediction profiler with the levels of the active factors all set to the middle level, and the two-factor interaction plot.

### 4.3 THE ONE-QUARTER FRACTION OF THE $2^k$ DESIGN

For a moderately large number of factors, smaller fractions of the  $2^k$  design are frequently useful. Consider a one-quarter fraction of the  $2^k$  design. This design contains  $2^{k-2}$  runs and is usually called a  $2^{k-2}$  fractional factorial.

The  $2^{k-2}$  design may be constructed by first writing down a basic design consisting of the runs associated with a full factorial in  $k - 2$  factors and then associating the two additional columns with generators that are appropriately chosen interactions involving the first  $k - 2$  factors. Thus, a one-quarter fraction of the  $2^k$  design has two generators. If  $P$  and  $Q$  represent the generators chosen, then  $I = P$  and  $I = Q$  are called **generating relations** for the design. The signs of  $P$  and  $Q$  (either  $+$  or  $-$ ) determine which one of the one-quarter fractions is produced. All four fractions associated with the choice of generators  $\pm P$  and  $\pm Q$  are members of the same family. The fraction for which both  $P$  and  $Q$  are positive is the **principal fraction**.

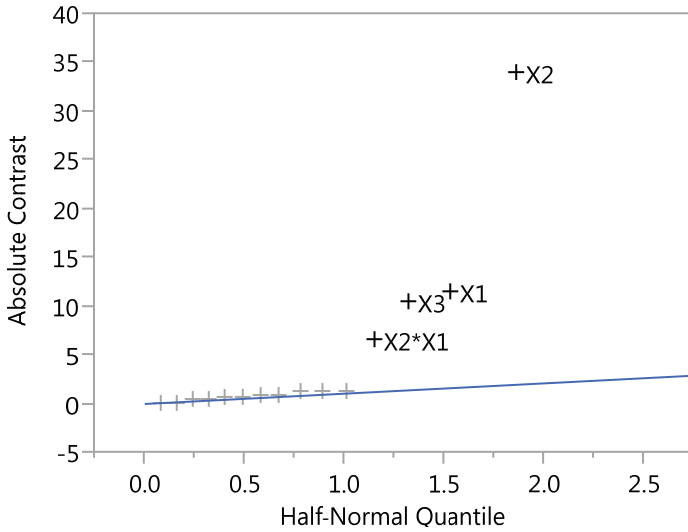


**TABLE 4.7 JMP Screening Output for Example 4.2**

**Screening for Y  
Contrasts**

Term	Contrast	Plot of $t$ -Ratio	Lenth $t$ -Ratio	Individual $P$ -Value	Simultaneous $P$ -Value
X2	34.0313		32.12	<0.0001*	<0.0001*
X1	11.6063		10.96	0.0003*	0.0012*
X3	10.4563		9.87	0.0004*	0.0018*
X4	-0.7063		-0.67	0.5002	1.0000
X5	0.1937		0.18	0.8682	1.0000
X2*X1	6.5813		6.21	0.0011*	0.0111*
X2*X3	0.4813		0.45	0.6788	1.0000
X1*X3	0.7063		0.67	0.5002	1.0000
X2*X4	1.3438		1.27	0.1938	0.9289
X1*X4	1.2938		1.22	0.2071	0.9480
X3*X4	0.4438		0.42	0.7015	1.0000
X2*X5	-0.1563		-0.15	0.8915	1.0000
X1*X5	1.2938		1.22	0.2071	0.9480
X3*X5	0.9438		0.89	0.3481	0.9993
X4*X5	-0.8688		-0.82	0.3861	1.0000

**Half-Normal Plot**



Lenth PSE = 1.05938

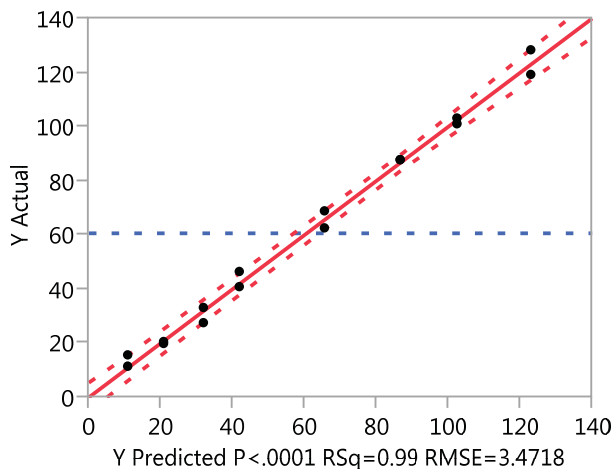
$P$ -Values derived from a simulation of 10,000 Lenth  $t$ -ratios.

**TABLE 4.8 Model Fitting Results for Example 4.2**

**Response Y**

**Whole Model**

**Actual by Predicted Plot**



**Summary of Fit**

RSquare	0.9943
RSquare Adj	0.992227
Root Mean Square Error	3.471794
Mean of Response	60.65625
Observations (or Sum Wgts)	16

**Analysis of Variance**

Source	DF	Sum of Squares	Mean Square	F-Ratio
Model	4	23127.633	5781.91	479.6930
Error	11	132.587	12.05	<b>Prob &gt; F</b>
C. Total	15	23260.219		<.0001*

**Lack of Fit**

Source	DF	Sum of Squares	Mean Square	F-Ratio
Lack of Fit	3	23.76187	7.9206	0.5823
Pure Error	8	108.82500	13.6031	<b>Prob &gt; F</b>
Total Error	11	132.58687		0.6431

**Max RSq**

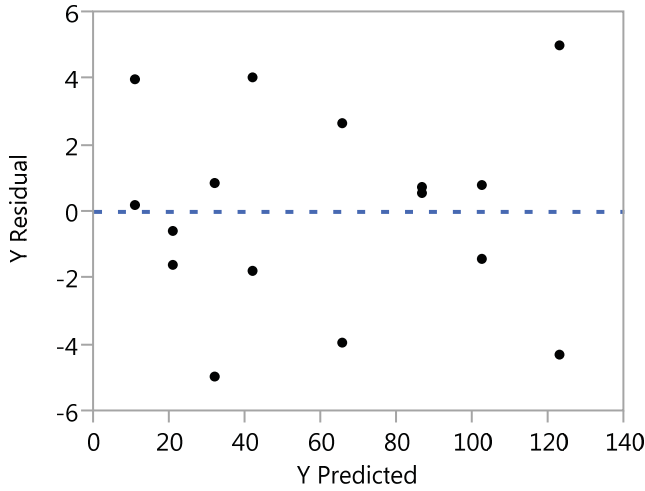
**Parameter Estimates**

Term	Estimate	Std Error	t-Ratio	Prob >  t
Intercept	60.65625	0.867948	69.88	<.0001*
X2	34.03125	0.867948	39.21	<.0001*
X1	11.60625	0.867948	13.37	<.0001*
X3	10.45625	0.867948	12.05	<.0001*
X2*X1	6.58125	0.867948	7.58	<.0001*

**TABLE 4.8 (Continued)**

Effect Tests					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
X2	1	1	18530.016	1537.333	<.0001*
X1	1	1	2155.281	178.8117	<.0001*
X3	1	1	1749.331	145.1323	<.0001*
X2*X1	1	1	693.006	57.4948	<.0001*

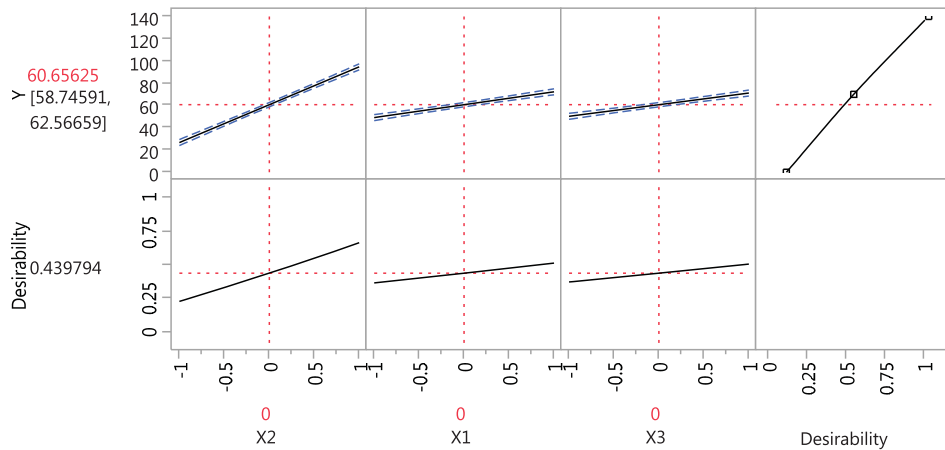
**Residual by Predicted Plot**



**Prediction Expression**

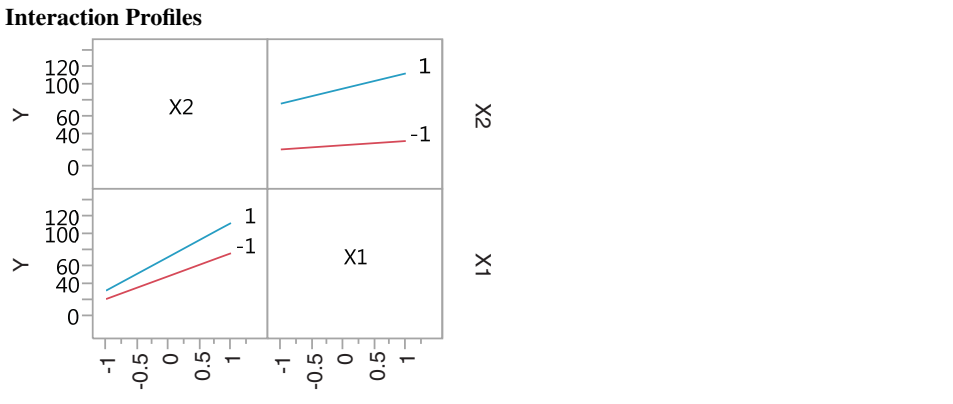
$$60.65625 + 34.03125 * X2 + 11.60625 * X1 + 10.45625 * X3 + X2 * [X1 * 6.58125]$$

**Prediction Profiler**



(continued)

TABLE 4.8 (Continued)



The **complete defining relation** for the design consists of all the columns that are equal to the identity column  $I$ . These will consist of  $P$ ,  $Q$ , and their **generalized interaction**  $PQ$ ; that is, the defining relation is  $I = P = Q = PQ$ . We call the elements  $P$ ,  $Q$ , and  $PQ$  in the defining relation **words**. The aliases of any effect are produced by the multiplication of the column for that effect by each word in the defining relation. Clearly, each effect has three aliases. The experimenter should be careful in choosing the generators so that potentially important effects are not aliased with each other.

As an example, consider the  $2^{6-2}$  design. Suppose we choose  $I = ABCE$  and  $I = BCDF$  as the design generators. Now the generalized interaction of the generators  $ABCE$  and  $BCDF$  is  $ADEF$ ; therefore, the complete defining relation for this design is

$$I = ABCE = BCDF = ADEF$$

Consequently, this design is of resolution IV. To find the aliases of any effect (e.g.,  $A$ ), multiply that effect by each word in the defining relation. For  $A$ , this produces

$$A = BCE = ABCDF = DEF$$

It is easy to verify that every main effect is aliased by three-factor and five-factor interactions, whereas two-factor interactions are aliased with each other and with higher-order interactions. Thus, when we estimate  $A$ , for example, we are really estimating  $A + BCE + DEF + ABCDF$ . The complete alias structure of this design is shown in Table 4.9. If three-factor and higher interactions are negligible, this design gives clear estimates of the main effects.

To construct the design, first write down the basic design, which consists of the 16 runs for a full  $2^{6-2} = 2^4$  design in  $A$ ,  $B$ ,  $C$ , and  $D$ . Then the two factors  $E$  and  $F$  are added by associating their plus and minus levels with the plus and minus signs of the interactions  $ABC$  and  $BCD$ , respectively. This procedure is shown in Table 4.10.

There are, of course, three alternate fractions of this particular  $2^{6-2}_{IV}$  design. They are the fractions with generating relationships  $I = ABCE$  and  $I = -BCDF$ ;  $I = -ABCE$  and  $I = BCDF$ ; and  $I = -ABCE$  and  $I = -BCDF$ . These fractions may be easily constructed by the method shown in Table 4.10. For example, if we wish to find the fraction for which

**TABLE 4.9** Alias Structure for the  $2^{6-2}_{IV}$  Design with  $I = ABCE = BCDF = ADEF$

$A = BCE = DEF = ABCDF$	$AB = CE = ACDF = BDEF$
$B = ACE = CDF = ABDEF$	$AC = BE = ABDF = CDEF$
$C = ABE = BDF = ACDEF$	$AD = EF = ABDF = CDEF$
$D = BCF = AEF = ABCDE$	$AE = BC = DF = ABCDEF$
$E = ABC = ADF = BCDEF$	$AF = DE = BCEF = ABCD$
$F = BCD = ADE = ABCEF$	$BD = CF = ACDE = ABEF$
	$BF = CD = ACEF = ABDE$
<hr/>	
$ABD = CDE = ACF = BEF$	
$ACD = BDE = ABF = CEF$	

$I = ABCE$  and  $I = -BCDF$ , then in the last column of Table 4.8 we set  $F = -BCD$ , and the column of levels for factor  $F$  becomes

+ + - - - - + + - - + + + + - -

The complete defining relation for this alternative fraction is  $I = ABCE = -BCDF = -ADEF$ . Certain signs in the alias structure in Table 4.10 are now changed; for instance, the aliases of  $A$  are  $A = BCE = -DEF = -ABCDF$ . Thus, the linear combination of the observations  $[A]$  actually estimates  $A + BCE - DEF - ABCDF$ .

Finally, note that the  $2^{6-2}_{IV}$  fractional factorial collapses to a single replicate of a  $2^4$  design in any subset of four factors that is not a word in the defining relation. It also collapses to a replicated one-half fraction of a  $2^4$  in any subset of four factors that is a word in the defining relation. Thus, the design in Table 4.10 becomes two replicates of a  $2^{4-1}$  in the factors  $ABCE$ ,  $BCDF$ , and  $ABEF$ , because these are the words in the defining relation. There are 12 other combinations of the six factors, such as  $ABCD$ ,  $ABDF$ , and so on, for

**TABLE 4.10** Construction of the  $2^{6-2}_{IV}$  Design with the Generators  $I = ABCE$  and  $I = BCDF$

| Run | $A$ | $B$ | $C$ | $D$ | $E = ABC$ | $F = BCD$ |
|-----|-----|-----|-----|-----|-----------|-----------|
| 1   | -   | -   | -   | -   | -         | -         |
| 2   | +   | -   | -   | -   | +         | -         |
| 3   | -   | +   | -   | -   | +         | +         |
| 4   | +   | +   | -   | -   | -         | +         |
| 5   | -   | -   | +   | -   | +         | +         |
| 6   | +   | -   | +   | -   | -         | +         |
| 7   | -   | +   | +   | -   | -         | -         |
| 8   | +   | +   | +   | -   | +         | -         |
| 9   | -   | -   | -   | +   | -         | +         |
| 10  | +   | -   | -   | +   | +         | +         |
| 11  | -   | +   | -   | +   | +         | -         |
| 12  | +   | +   | -   | +   | -         | -         |
| 13  | -   | -   | +   | +   | +         | -         |
| 14  | +   | -   | +   | +   | -         | -         |
| 15  | -   | +   | +   | +   | -         | +         |
| 16  | +   | +   | +   | +   | +         | +         |

which the design projects to a single replicate of the  $2^4$ . This design also collapses to two replicates of a  $2^3$  in any subset of three of the six factors or four replicates of a  $2^2$  in any subset of two factors.

In general, any  $2^{k-2}$  fractional factorial design can be collapsed into either a full factorial or a fractional factorial in some subset of  $r \leq k - 2$  of the original factors. Those subsets of variables that form full factorials in  $r = k - 2$  are not words in the complete defining relation. Any subset of  $r = 3$  factors will give a full factorial design.

**Example 4.3 The Injection Molding Experiment** Parts manufactured in an injection molding process are showing excessive shrinkage. This is causing problems in assembly operations downstream from the injection molding area. A quality improvement team has decided to use a designed experiment to study the injection molding process so that shrinkage can be reduced. The team decides to investigate six factors—mold temperature ( $A$ ), screw speed ( $B$ ), holding time ( $C$ ), cycle time ( $D$ ), gate size ( $E$ ), and holding pressure ( $F$ )—each at two levels, with the objective of learning how each factor affects shrinkage and also something about how the factors interact.

The team decides to use the 16-run two-level fractional factorial design in Table 4.11 along with four center points, so that the linearity of the response function can be investigated via a curvature test. The design is shown in Table 4.11, along with the observed shrinkage ( $\times 10$ ) for the test part produced at each of the 20 runs in the design.

A normal probability plot of the effect estimates from this experiment is shown in Fig. 4.11. The only large effects are  $A$  (mold temperature),  $B$  (screw speed), and the  $AB$  interaction. In light of the alias relationships in Table 4.9, it seems reasonable to adopt these conclusions tentatively. Table 4.12 shows the analysis of variance for this experiment.

**TABLE 4.11** A  $2^{6-2}_{IV}$  Design for the Injection Molding Experiment in Example 4.3

| Run | $A$ | $B$ | $C$ | $D$ | $E = ABC$ | $F = BCD$ | Observed Shrinkage ( $\times 10$ ) |
|-----|-----|-----|-----|-----|-----------|-----------|------------------------------------|
| 1   | –   | –   | –   | –   | –         | –         | 6                                  |
| 2   | +   | –   | –   | –   | +         | –         | 10                                 |
| 3   | –   | +   | –   | –   | +         | +         | 32                                 |
| 4   | +   | +   | –   | –   | –         | +         | 60                                 |
| 5   | –   | –   | +   | –   | +         | +         | 4                                  |
| 6   | +   | –   | +   | –   | –         | +         | 15                                 |
| 7   | –   | +   | +   | –   | –         | –         | 26                                 |
| 8   | +   | +   | +   | –   | +         | –         | 60                                 |
| 9   | –   | –   | –   | +   | –         | +         | 8                                  |
| 10  | +   | –   | –   | +   | +         | +         | 12                                 |
| 11  | –   | +   | –   | +   | +         | –         | 34                                 |
| 12  | +   | +   | –   | +   | –         | –         | 60                                 |
| 13  | –   | –   | +   | +   | +         | –         | 16                                 |
| 14  | +   | –   | +   | +   | –         | –         | 5                                  |
| 15  | –   | +   | +   | +   | –         | +         | 37                                 |
| 16  | +   | +   | +   | +   | +         | +         | 52                                 |
| 17  | 0   | 0   | 0   | 0   | 0         | 0         | 29                                 |
| 18  | 0   | 0   | 0   | 0   | 0         | 0         | 34                                 |
| 19  | 0   | 0   | 0   | 0   | 0         | 0         | 26                                 |
| 20  | 0   | 0   | 0   | 0   | 0         | 0         | 30                                 |

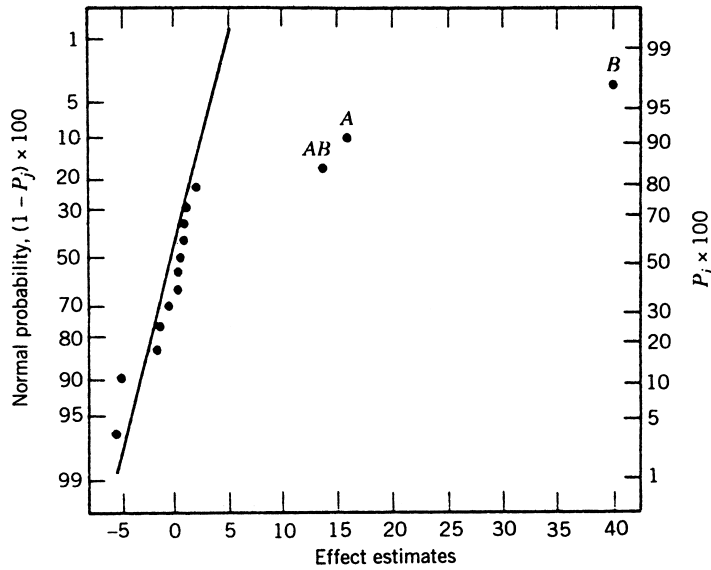


Figure 4.11 Normal probability plot of effects for Example 4.3.

The pure error sum of squares is computed from the replicate runs at the center, and the lack-of-fit sum of squares is formed by pooling all the effects that appeared small on the normal probability plot. Notice that the lack-of-fit (higher-order interaction) and curvature (quadratic) tests have small  $F$ -ratios, so the model is correctly specified.

The plot of the  $AB$  interaction in Fig. 4.12 shows that the process is insensitive to temperature if the screw speed is at the low level but very sensitive to temperature if the screw speed is at the high level. With the screw speed at the low level, the process should produce an average shrinkage of around 10% regardless of the temperature level chosen.

Based on this initial analysis, the team decides to set both the mold temperature and the screw speed at the low level. This set of conditions will reduce the mean shrinkage of parts to around 10%. However, the variability in shrinkage from part to part is still a potential problem. In effect, the mean shrinkage can be substantially reduced by the above modifications; however, the part-to-part variability in shrinkage over a production run could

TABLE 4.12 Analysis of Variance for the Injection Molding Experiment (Example 4.3)

| Source of Variation | Sum of Squares | Degrees of Freedom | Mean Squares | $F_0$   | $P$ -Value             |
|---------------------|----------------|--------------------|--------------|---------|------------------------|
| <i>A</i>            | 770.063        | 1                  | 770.063      | 41.033  | $1.18 \times 10^{-5}$  |
| <i>B</i>            | 5076.563       | 1                  | 5076.563     | 270.505 | $5.25 \times 10^{-11}$ |
| <i>AB</i>           | 564.063        | 1                  | 564.063      | 30.056  | 0.0001                 |
| Curvature           | 19.012         | 1                  | 19.012       | 1.742   | 0.2786                 |
| Residual            | 281.500        | 15                 | 18.767       |         |                        |
| Lack-of-fit         | (248.750)      | (12)               | 20.729       | 1.899   | 0.3277                 |
| Pure error          | (32.750)       | (3)                | 10.917       |         |                        |
| Total               | 6711.200       | 19                 |              |         |                        |

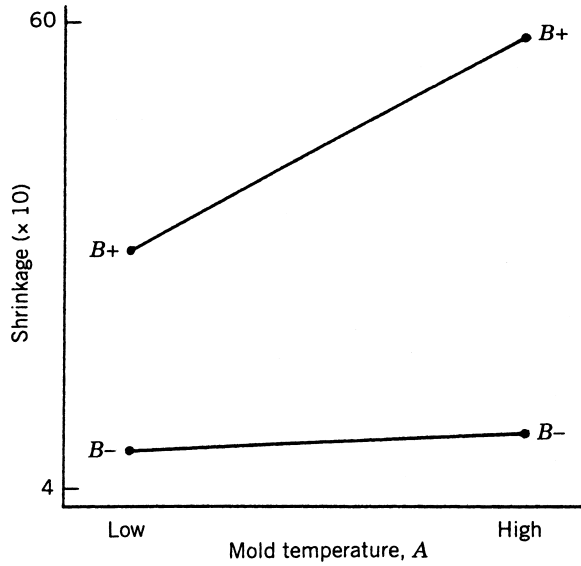


Figure 4.12 Plot of AB (mold-temperature–screw-speed) interaction for Example 4.3.

still cause problems in assembly. One way to address this issue is to see if any of the process factors affect the variability in parts shrinkage.

Figure 4.13 presents the normal probability plot of the residuals. This plot appears satisfactory. The plots of residuals versus each factor were then constructed. One of these plots, that for residuals versus factor C (holding time), is shown in Fig. 4.14. The plot reveals that there is much less scatter in the residuals at the low holding time than at the

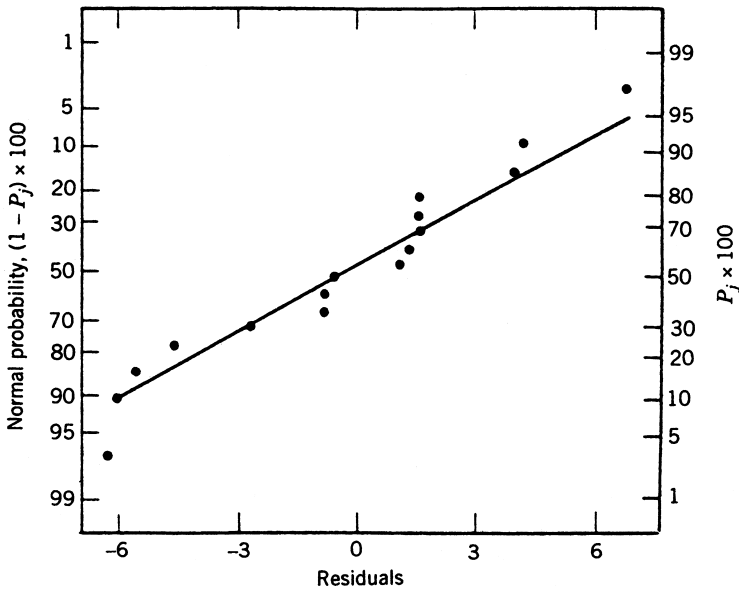
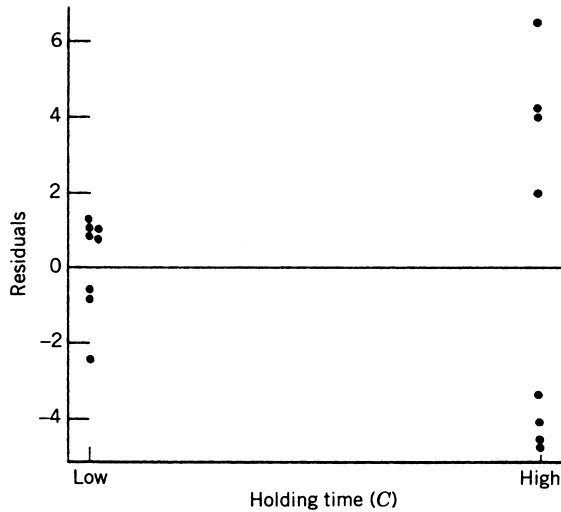


Figure 4.13 Normal probability plot of residuals for Example 4.3.





**Figure 4.14** Residuals versus holding time (C) for Example 4.4.

high holding time. These residuals were obtained in the usual way from the model for predicted shrinkage

$$\begin{aligned}\hat{y} &= b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 \\ &= 27.3125 + 6.9375x_1 + 17.8125x_2 + 5.9375x_1x_2\end{aligned}$$

where  $x_1$ ,  $x_2$ , and  $x_1x_2$  are the coded variables that correspond to the factors  $A$  and  $B$  and the  $AB$  interaction. The residuals are then

$$e = y - \hat{y}$$

The regression model used to produce the residuals essentially removes the location effects of  $A$ ,  $B$ , and  $AB$  from the data; the residuals therefore contain information about unexplained variability. Figure 4.14 indicates that there is a pattern in the variability and that the variability in shrinkage of parts may be smaller when the holding time is at the low level.

Figure 4.15 shows the data from this experiment projected onto a cube in the factors  $A$ ,  $B$ , and  $C$ . The average observed shrinkage and the range of observed shrinkage are shown at each corner of the cube. From inspection of this figure, we see that running the process with the screw speed ( $B$ ) at the low level is the key to reducing average parts shrinkage. If  $B$  is low, virtually any combination of temperature ( $A$ ) and holding time ( $C$ ) will result in low values of average part shrinkage. However, from examining the ranges of the shrinkage values at each corner of the cube, it is immediately clear that setting the holding time ( $C$ ) at the low level is the only reasonable choice if we wish to keep the part-to-part variability in shrinkage low during a production run.

In many response surface problems it is useful to actually build a model for both the *mean response* and a measure of the *variability* in response (such as the range, the variance, or the standard deviation). In fact, we illustrated this in Example 3.3. Then the optimization problem is a tradeoff between these two criteria, such as minimizing the variability while simultaneously moving the mean response to a desired or target value. We will present several techniques for doing this in subsequent chapters.

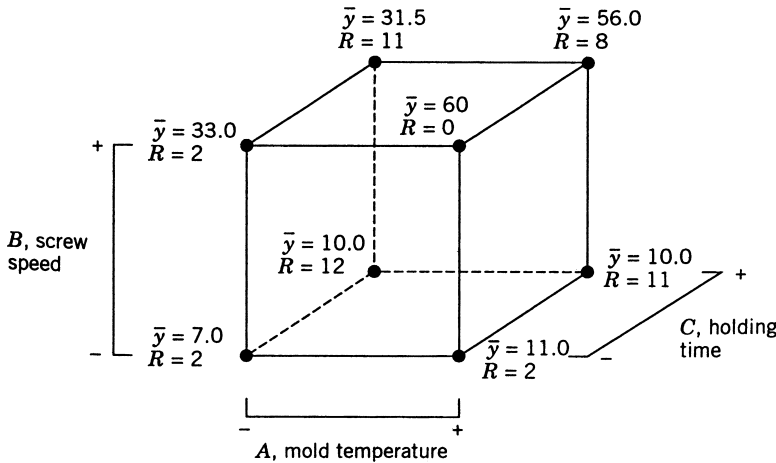


Figure 4.15 Average shrinkage ( $\bar{y}$ ) and range of shrinkage ( $R$ ) in factors A, B, and C for Example 4.3.

#### 4.4 THE GENERAL $2^{k-p}$ FRACTIONAL FACTORIAL DESIGN

A  $2^k$  fractional factorial design containing  $2^{k-p}$  runs is called a  $1/2^p$  fraction of the  $2^k$  design or, more simply, a  $2^{k-p}$  fractional factorial design. These designs require the selection of  $p$  independent design generators. The defining relation for the design consists of the  $p$  generators initially chosen and their  $2^p - p - 1$  generalized interactions. In this section we discuss the construction and analysis of these designs.

The alias structure may be found by multiplying each effect column by the defining relations. Care should be exercised in choosing the generators so that effects of potential interest are not aliased with each other. Each effect has  $2^p - 1$  aliases. For moderately large values of  $k$ , we usually assume higher-order interactions (say, third- or fourth-order and higher) to be negligible, and this greatly simplifies the alias structure.

It is important to select the  $p$  generators for a  $2^{k-p}$  fractional factorial design in such a way that we obtain the best possible alias relationships. A reasonable criterion is to select the generators such that the resulting  $2^{k-p}$  design has the highest possible resolution. Table 4.13 presents a selection of  $2^{k-p}$  fractional factorial designs for  $k \leq 11$  factors and up to  $n \leq 128$  runs, adapted from Montgomery (2013). The suggested generators in this table will result in a design of the highest possible resolution. These designs are also **minimum aberration** designs; that is if the design has resolution  $R$ , then it has the minimum number of main effects aliased with interactions of order  $R - 1$ , the minimum number of two-factor interactions aliased with interactions of order  $R - 2$ , and so on. The alias relationships for all of the designs in Table 4.13 for which  $n \leq 64$  are given in Appendix Table XII(a–v) of Montgomery (2013).

**Example 4.4 Selection of a Fractional Factorial** To illustrate the use of Table 4.13, suppose that we have seven factors and that we are interested in estimating the seven main effects and getting some insight regarding the two-factor interactions. We are willing to assume that three-factor and higher interactions are negligible. This information suggests that a resolution IV design would be appropriate.

**TABLE 4.13 Selected  $2^{k-p}$  Fractional Factorial Designs**

| Number of Factors, $k$ | Fraction        | Number of Runs | Design Generators  |
|------------------------|-----------------|----------------|--|
| 3                      | $2^{3-1}_{III}$ | 4              | $C = \pm AB$   |
| 4                      | $2^{4-1}_{IV}$  | 8              | $D = \pm ABC$  |
| 5                      | $2^{5-1}_V$     | 16             | $E = \pm ABCD$   |
|                        | $2^{5-2}_{III}$ | 8              | $D = \pm AB$<br>$E = \pm AC$   |
| 6                      | $2^{6-1}_{VI}$  | 32             | $F = \pm ABCDE$  |
|                        | $2^{6-2}_{IV}$  | 16             | $E = \pm ABC$<br>$F = \pm BCD$   |
|                        | $2^{6-3}_{III}$ | 8              | $D = \pm AB$<br>$E = \pm AC$<br>$F = \pm BC$                                       |
| 7                      | $2^{7-1}_{VII}$ | 64             | $G = \pm ABCDEF$   |
|                        | $2^{7-2}_{IV}$  | 32             | $F = \pm ABCD$<br>$G = \pm ABDE$   |
|                        | $2^{7-3}_{IV}$  | 16             | $E = \pm ABC$<br>$F = \pm BCD$<br>$G = \pm ACD$                                    |
|                        | $2^{7-4}_{III}$ | 8              | $D = \pm AB$<br>$E = \pm AC$<br>$F = \pm BC$<br>$G = \pm ABC$                      |
| 8                      | $2^{8-2}_V$     | 64             | $G = \pm ABCD$<br>$H = \pm ABEF$   |
|                        | $2^{8-3}_{IV}$  | 32             | $F = \pm ABC$<br>$G = \pm ABD$<br>$H = \pm BCDE$                                   |
|                        | $2^{8-4}_{IV}$  | 16             | $E = \pm BCD$<br>$F = \pm ACD$<br>$G = \pm ABC$<br>$H = \pm ABD$                   |
| 9                      | $2^{9-2}_{VI}$  | 128            | $H = \pm ACDFG$<br>$J = \pm BCEFG$   |
|                        | $2^{9-3}_{IV}$  | 64             | $G = \pm ABCD$<br>$H = \pm ACEF$<br>$J = \pm CDEF$                                 |
|                        | $2^{9-4}_{IV}$  | 32             | $F = \pm BCDE$<br>$G = \pm ACDE$<br>$H = \pm ABDE$<br>$J = \pm ABCE$               |
|                        | $2^{9-5}_{III}$ | 16             | $E = \pm ABC$<br>$F = \pm BCD$<br>$G = \pm ACD$<br>$H = \pm ABD$<br>$J = \pm ABCD$ |

(continued)

TABLE 4.13 (Continued)

| Number of Factors, $k$ | Fraction                | Number of Runs | Design Generators  |
|------------------------|-------------------------|----------------|--|
| 10                     | $2^{10-3}_{\text{V}}$   | 128            | $H = \pm ABCG$<br>$J = \pm ACDE$<br>$K = \pm ACDF$   |
|                        | $2^{10-4}_{\text{IV}}$  | 64             | $G = \pm BCDF$<br>$H = \pm ACDF$<br>$J = \pm ABDE$<br>$K = \pm ABCE$   |
|                        | $2^{10-5}_{\text{IV}}$  | 32             | $F = \pm ABCD$<br>$G = \pm ABCE$<br>$H = \pm ABDE$<br>$J = \pm ACDE$<br>$K = \pm BCDE$                             |
|                        | $2^{10-6}_{\text{III}}$ | 16             | $E = \pm ABC$<br>$F = \pm BCD$<br>$G = \pm ACD$<br>$H = \pm ABD$<br>$J = \pm ABCD$<br>$K = \pm AB$                 |
| 11                     | $2^{11-5}_{\text{IV}}$  | 64             | $G = \pm CDE$<br>$H = \pm ABCD$<br>$J = \pm ABF$<br>$K = \pm BDEF$<br>$L = \pm ADEF$                               |
|                        | $2^{11-6}_{\text{IV}}$  | 32             | $F = \pm ABC$<br>$G = \pm BCD$<br>$H = \pm CDE$<br>$J = \pm ACD$<br>$K = \pm ADE$<br>$L = \pm BDE$                 |
|                        | $2^{11-7}_{\text{III}}$ | 16             | $E = \pm ABC$<br>$F = \pm BCD$<br>$G = \pm ACD$<br>$H = \pm ABD$<br>$J = \pm ABCD$<br>$K = \pm AB$<br>$L = \pm AC$ |

Table 4.13 shows that there are two resolutions IV fractions available: the  $2^{7-2}_{\text{IV}}$  design with 32 runs and the  $2^{7-3}_{\text{IV}}$  with 16 runs. Consider the 16-run design. This is a one-eighth fraction with generators  $E = \pm ABC$ ,  $F = \pm BCD$ , and  $G = \pm ACD$ , and if we choose the principal fraction the complete defining relation is

$$I = ABCE = BCDF = ADEF = ACDG = BDEG = ABFG = CEF G$$

Now because the design is of resolution IV, it is clear that the main effects will be aliased with at worst a three-factor interaction. So if we can safely ignore these higher-order

interactions, this design will give excellent information on the main effects. The two-factor interaction alias chains are

$$\begin{aligned}
 AB &= CE = FG \\
 AC &= BE = DG \\
 AD &= EF = CG \\
 AE &= BC = DF \\
 AF &= DE = BG \\
 AG &= CD = BF \\
 BD &= CF = EG
 \end{aligned}$$

when ignoring three-factor and higher interactions. Thus this design would likely be satisfactory to satisfy the experimenter’s objectives.

The complete design is shown in Table 4.14. Notice that it was constructed by starting with the 16-run  $2^4$  design in  $A, B, C,$  and  $D$  as the basic design and then adding the three columns  $E = ABC, F = BCD,$  and  $G = ACD$  as suggested in Table 4.13.

**Projection of the  $2^{k-p}$  Fractional Factorial** The  $2^{k-p}$  design collapses into either a full factorial or a fractional factorial in any subset of  $r \leq k - p$  of the original factors. Those subsets of factors providing fractional factorials are subsets appearing as words in the complete defining relation. Furthermore, a design of resolution  $R$  will project into a full factorial in any subset of  $R - 1$  factors. These results are particularly useful in screening experiments when we suspect at the outset of the experiment that most of the original factors will have small effects. The original  $2^{k-p}$  fractional factorial can then be projected

**TABLE 4.14** A  $2_{IV}^{7-3}$  Fractional Factorial Design

| Run | Basic Design |     |     |     | $E = ABC$ | $F = BCD$ | $G = ACD$ |
|-----|--------------|-----|-----|-----|-----------|-----------|-----------|
|     | $A$          | $B$ | $C$ | $D$ |           |           |           |
| 1   | -            | -   | -   | -   | -         | -         | -         |
| 2   | +            | -   | -   | -   | +         | -         | +         |
| 3   | -            | +   | -   | -   | +         | +         | -         |
| 4   | +            | +   | -   | -   | -         | +         | +         |
| 5   | -            | -   | +   | -   | +         | +         | +         |
| 6   | +            | -   | +   | -   | -         | +         | -         |
| 7   | -            | +   | +   | -   | -         | -         | +         |
| 8   | +            | +   | +   | -   | +         | -         | -         |
| 9   | -            | -   | -   | +   | -         | +         | +         |
| 10  | +            | -   | -   | +   | +         | +         | -         |
| 11  | -            | +   | -   | +   | +         | -         | +         |
| 12  | +            | +   | -   | +   | -         | -         | -         |
| 13  | -            | -   | +   | +   | +         | -         | -         |
| 14  | +            | -   | +   | +   | -         | -         | +         |
| 15  | -            | +   | +   | +   | -         | +         | -         |
| 16  | +            | +   | +   | +   | +         | +         | +         |

into a full factorial, say, in the most interesting factors. Conclusions drawn from designs of this type should be considered tentative and subject to further analysis. It is usually possible to find alternative explanations of the data involving higher-order interactions.

As an example, consider the  $2_{IV}^{7-3}$  design from Example 4.4. This is a 16-run design involving seven factors. It will project into a full factorial in any four of the original seven factors that is not a word in the complete defining relation (see Table 4.14). Thus, there are 28 subsets of four factors that would form  $2^4$  designs. One combination that is obvious upon inspecting Table 4.14 is  $A, B, C,$  and  $D$ . It will also form a replicated full factorial in any subset of three of the original seven factors.

To illustrate the usefulness of this projection properly, suppose that we are conducting an experiment to improve the performance of a chemical process and the seven factors are:

1. Concentration of reactant  $A$
2. Temperature
3. Feed rate
4. Time
5. Agitation rate
6. Concentration of reactant  $B$
7. Catalyst type

We are fairly certain that time, temperature, catalyst type, and concentration of reactant  $A$  will affect performance and that these factors may interact. The role of the other three factors is less well known, but it is likely that they are negligible. A reasonable strategy would be to assign these four factors to columns  $A, B, C,$  and  $D$ , respectively, in Table 4.14. The remaining three factors would be assigned to columns  $E, F,$  and  $G$ , respectively. If we are correct and the “minor variables”  $E, F,$  and  $G$  are negligible, we will be left with a full  $2^k$  design in the key process variables.

#### 4.5 RESOLUTION III DESIGNS

As indicated earlier, the sequential use of fractional factorial designs is very useful, often leading to great economy and efficiency in experimentation. We now illustrate these ideas using the class of resolution III designs. Box and Hunter (1961a,b) are excellent references for the technical details.

It is possible to construct resolution III designs for investigating up to  $k = N - 1$  factors in only  $N$  runs, where  $N$  is a multiple of 4. These designs are frequently useful in industrial experimentation. Designs in which  $N$  is a power of 2 can be constructed by the methods presented earlier in this chapter, and these are presented first. Of particular importance are designs requiring 4 runs for up to 3 factors, 8 runs for up to 7 factors, and 16 runs for up to 15 factors. If  $k = N - 1$ , the fractional factorial design is said to be **saturated**.

A design for analyzing up to three factors in four runs is the  $2_{III}^{3-1}$  design, presented in Section 4.2. Another very useful saturated fractional factorial is a design for studying up to seven factors in eight runs—that is, the  $2_{III}^{7-4}$  design. This design is a one-sixteenth fraction of the  $2^7$ . It may be constructed by first writing down as the basic design the plus and minus levels for a full  $2^3$  design in  $A, B,$  and  $C$  and then associating the levels of four additional factors with the interactions of the original three as follows:  $D = AB, E = AC, F = BC,$  and

**TABLE 4.15** The Saturated  $2_{III}^{7-4}$  Design with the Generators  $I = ABD, I = ACE, I = BCF,$  and  $I = ABCG$ 

| Run | A | B | C | D = AB | E = AC | F = BC | G = ABC |                |
|-----|---|---|---|--------|--------|--------|---------|----------------|
| 1   | - | - | - | +      | +      | +      | -       | <i>def</i>     |
| 2   | + | - | - | -      | -      | +      | +       | <i>afg</i>     |
| 3   | - | + | - | -      | +      | -      | +       | <i>beg</i>     |
| 4   | + | + | - | +      | -      | -      | -       | <i>abd</i>     |
| 5   | - | - | + | +      | -      | -      | +       | <i>cdg</i>     |
| 6   | + | - | + | -      | +      | -      | -       | <i>ace</i>     |
| 7   | - | + | + | -      | -      | +      | -       | <i>cdg</i>     |
| 8   | + | + | + | +      | +      | +      | +       | <i>abcdefg</i> |

$G = ABC$ . Thus, the generators for this design are  $I = ABD, I = ACE, I = BCF,$  and  $I = ABCG$ . The design is shown in Table 4.15.

The complete defining relation for this design is obtained by multiplying the four generators  $ABD, ACE, BCF,$  and  $ABCG$  together two at a time, three at a time, and four at a time, yielding

$$\begin{aligned}
 I &= ABD = ACE = BCF = ABCG = BCDE = ACDF = CDG \\
 &= ABEF = BEG = AFG = DEF = ADEG = CEFG = BDFG \\
 &= ABCDEFG
 \end{aligned}$$

To find the aliases of any effect, simply multiply the effect by each word in the defining relation. For example, the aliases of  $B$  are

$$\begin{aligned}
 B &= AD = ABCE = CF = ACG = CDE = ABCDF = BCDG = AEF = EG \\
 &= ABFG = BDEF = ABDEG = BCEFG = DFG = ACDEFG
 \end{aligned}$$

This design is a one-sixteenth fraction; and because the signs chosen for the generators are positive, this is the principal fraction. It is also of resolution III, because the smallest number of letters in any word of the defining contrast is three. Any one of the 16 different  $2_{III}^{7-4}$  designs in this family could be constructed by using the generators with one of the 16 possible arrangements of signs in  $I = \pm ABC, I = \pm ACE, I = \pm BCF, I = \pm ABCG$ .

The seven degrees of freedom in this design may be used to estimate the seven main effects. Each of these effects has 15 aliases; however, if we assume that three-factor and higher interactions are negligible, then considerable simplification in the alias structure results. Making this assumption, each of the linear combinations associated with the seven main effects in this design actually estimates the main effect and three two-factor interactions:

$$\begin{aligned}
 [A] &\longrightarrow A + BD + CE + FG \\
 [B] &\longrightarrow B + AD + CF + EG \\
 [C] &\longrightarrow C + AE + BF + DG \\
 [D] &\longrightarrow D + AB + CG + EF \\
 [E] &\longrightarrow E + AC + BG + DF \\
 [F] &\longrightarrow F + BC + AG + DE \\
 [G] &\longrightarrow G + CD + BE + AF
 \end{aligned} \tag{4.1}$$

**TABLE 4.16** A  $2^{6-3}_{III}$  Design with the Generators  $I = ABD, I = ACE, \text{ and } I = BCF$

| Run | A | B | C | D = AB | E = AC | F = BC |               |
|-----|---|---|---|--------|--------|--------|---------------|
| 1   | - | - | - | +      | +      | +      | <i>def</i>    |
| 2   | + | - | - | -      | -      | +      | <i>af</i>     |
| 3   | - | + | - | -      | +      | -      | <i>be</i>     |
| 4   | + | + | - | +      | -      | -      | <i>abd</i>    |
| 5   | - | - | + | +      | -      | -      | <i>cd</i>     |
| 6   | + | - | + | -      | +      | -      | <i>ace</i>    |
| 7   | - | + | + | -      | -      | +      | <i>bcf</i>    |
| 8   | + | + | + | +      | +      | +      | <i>abcdef</i> |

In obtaining these aliases, we have ignored the three-factor and higher-order interactions, assuming that they will be negligible in most practical applications where a design of this type would be considered.

The saturated  $2^{7-4}_{III}$  design in Table 4.15 can be used to obtain resolution III designs for studying fewer than seven factors in eight runs. For example, to generate a design for six factors in eight runs, simply drop any one column in Table 4.15, for example, column G. This produces the design shown in Table 4.16.

It is easy to verify that this design is also of resolution III; in fact, it is a  $2^{6-3}_{III}$ , or a one-eighth fraction of the  $2^6$  design. The defining relation for the  $2^{6-3}_{III}$  design is equal to the defining relation for the original  $2^{7-4}_{III}$  design with any words containing the letter G deleted. Thus, the defining relation for our new design is

$$I = ABD = ACE = BCF = BCDE = ACDF = ABEF = DEF$$

In general, when  $d$  factors are dropped to produce a new design, the new defining relation is obtained as those words in the original defining relation that do not contain any dropped letters. When constructing designs by this method, care should be exercised to obtain the best arrangement possible. If we drop columns B, D, F, and G from Table 4.13, we obtain a design for three factors in eight runs, yet the treatment combinations correspond to two replicates of a  $2^2$  design. The experimenter would probably prefer to run a full  $2^3$  design in A, C, and E.

It is also possible to obtain a resolution III design for studying up to 15 factors in 16 runs. This saturated  $2^{15-11}_{III}$  design can be generated by first writing down the 16 treatment combinations associated with a  $2^4$  design in A, B, C, and D and then equating 11 new factors with the two-, three-, and four-factor interactions of the original four. In this design, each of the 15 main effects is aliased with seven two-factor interactions. A similar procedure can be used for the  $2^{31-26}_{III}$  design, which allows up to 31 factors to be studied in 32 runs.

**Sequential Use of Fractions to Separate Effects** By combining fractional factorial designs in which certain signs are switched, we can systematically isolate effects of potential interest. The alias structure for any fraction with the signs for one or more factors reversed is obtained by making changes of sign on the appropriate factors in the alias structure of the original fraction. This general procedure is called **fold-over**, and it is used in resolution III designs to break the links between main effects and two-factor interactions. In a **full fold-over**, we add to a resolution III fractional a second fraction in which the



TABLE 4.17 A  $2_{III}^{7-4}$  Design for the Eye Focus Time Experiment

| Run | A | B | C | D = AB | E = AC | F = BC | G = ABC |                | Time  |
|-----|---|---|---|--------|--------|--------|---------|----------------|-------|
| 1   | - | - | - | +      | +      | +      | -       | <i>def</i>     | 85.5  |
| 2   | + | - | - | -      | -      | +      | +       | <i>afg</i>     | 75.1  |
| 3   | - | + | - | -      | +      | -      | +       | <i>beg</i>     | 93.2  |
| 4   | + | + | - | +      | -      | -      | -       | <i>abd</i>     | 145.4 |
| 5   | - | - | + | +      | -      | -      | +       | <i>cdg</i>     | 83.7  |
| 6   | + | - | + | -      | +      | -      | -       | <i>ace</i>     | 77.6  |
| 7   | - | + | + | -      | -      | +      | -       | <i>bcf</i>     | 95.0  |
| 8   | + | + | + | +      | +      | +      | +       | <i>abcdefg</i> | 141.8 |

signs for all the factors are reversed. We may now use the combined design to estimate all the main effects clear of any two-factor interactions. The following example illustrates the technique.

**Example 4.5 Fold-Over of a Fractional Factorial** A human performance analyst is conducting an experiment to study eye focus time and has built an apparatus in which several factors can be controlled during the test. The factors he initially regards as important are acuity or sharpness of vision (*A*), distance from target to eye (*B*), target shape (*C*), illumination level (*D*), target size (*E*), target density (*F*), and subject (*G*). Two levels of each factor are considered. He suspects that only a few of these seven factors are of major importance and that high-order interactions between the factors can be neglected. On the basis of this assumption, the analyst decides to run a screening experiment to identify the most important factors and then to concentrate further study on those. To screen these seven factors, he runs the treatment combinations from the  $2_{III}^{7-4}$  design in Table 4.15 in random order, obtaining the focus times in milliseconds, as shown in Table 4.17.

Seven main effects and their aliases may be estimated from these data. From Equation 4.1, we see that the effects and their aliases are

$$[A] = 20.63 \longrightarrow A + BD + CE + FG$$

$$[B] = 38.38 \longrightarrow B + AD + CF + EG$$

$$[C] = -0.28 \longrightarrow C + AE + BF + DG$$

$$[D] = 28.88 \longrightarrow D + AB + CG + EF$$

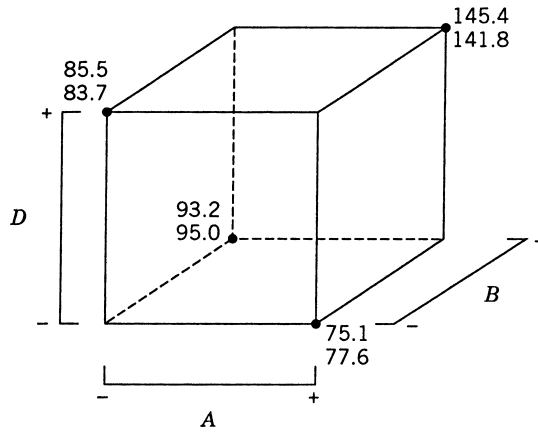
$$[E] = -0.28 \longrightarrow E + AC + BG + DF$$

$$[F] = -0.63 \longrightarrow F + BC + AG + DE$$

$$[G] = -2.43 \longrightarrow G + CD + BE + AF$$

For example,

$$\begin{aligned} [A] &= \frac{1}{4}(-85.5 + 75.1 - 93.2 + 145.4 - 83.7 + 77.6 - 95.0 + 141.8) \\ &= 20.63 \end{aligned}$$



**Figure 4.16** The  $2^{7-4}_{III}$  design projected into two replicates of a  $2^{3-1}_{III}$  design in  $A$ ,  $B$ , and  $D$ .

The largest three effects are  $[A]$ ,  $[B]$ , and  $[D]$ . The simplest interpretation of the data is that the main effects of  $A$ ,  $B$ , and  $D$  are all significant. However, this interpretation is not unique, because one could also logically conclude that  $A$ ,  $B$ , and the  $AB$  interaction, or perhaps  $B$ ,  $D$ , and the  $BD$  interaction, or perhaps  $A$ ,  $D$ , and the  $AD$  interaction are the true effects.

Notice that  $ABD$  is a word in the defining relation for this design. Therefore, this  $2^{7-4}_{III}$  design does not project into a full  $2^3$  factorial in  $A$ ,  $B$ , and  $D$ ; instead, it projects into two replicates of a  $2^{3-1}$  design, as shown in Fig. 4.16. Because the  $2^{3-1}$  design is a resolution III design,  $A$  will be aliased with  $BD$ ,  $B$  will be aliased with  $AD$ , and  $D$  will be aliased with  $AB$ , so the interactions cannot be separated from the main effects. The analyst here may have been unlucky. If he had assigned illumination level to column  $C$  in Table 4.15 instead of  $D$ , the design would have projected into a  $2^3$  design.

To separate the main effects and the two-factor interactions, a second fraction is run with all the signs reversed. This fold-over design is shown in Table 4.18 along with the observed responses. Notice that when we fold over a resolution III design in this manner, we (in effect) change the signs on the generators that have an odd number of letters. The effects estimated by this fraction are

$$[A]' = -17.68 \longrightarrow A - BD - CE - FG$$

$$[B]' = 37.73 \longrightarrow B - AD - CF - EG$$

**TABLE 4.18** Fold-over of the  $2^{7-4}_{III}$  Design in Table 4.15

| Run | $A$ | $B$ | $C$ | $D = -AB$ | $E = -AC$ | $F = -BC$ | $G = ABC$ |        | Time  |
|-----|-----|-----|-----|-----------|-----------|-----------|-----------|--------|-------|
| 1   | +   | +   | +   | -         | -         | -         | +         | $abcg$ | 91.3  |
| 2   | -   | +   | +   | +         | +         | -         | -         | $bcde$ | 136.7 |
| 3   | +   | -   | +   | +         | -         | +         | -         | $acdf$ | 82.4  |
| 4   | -   | -   | +   | -         | +         | +         | +         | $cefg$ | 73.4  |
| 5   | +   | +   | -   | -         | +         | +         | -         | $abef$ | 94.1  |
| 6   | -   | +   | -   | +         | -         | +         | +         | $bdfg$ | 143.8 |
| 7   | +   | -   | -   | +         | +         | -         | +         | $adeg$ | 87.3  |
| 8   | -   | -   | -   | -         | -         | -         | -         | (1)    | 71.9  |

$$\begin{aligned}
 [C]' &= -3.33 \longrightarrow C - AE - BF - DG \\
 [D]' &= 29.88 \longrightarrow D - AB - CG - EF \\
 [E]' &= 0.53 \longrightarrow E - AC - BG - DF \\
 [F]' &= 1.63 \longrightarrow F - BC - AG - DE \\
 [G]' &= 2.68 \longrightarrow G - CD - BE - AF
 \end{aligned}$$

By combining the effect estimates from this second fraction with the effect estimates from the original eight runs, we obtain the following estimates of the effects:

| <i>i</i> | From $\frac{1}{2}([i] + [i]')$ | From $\frac{1}{2}([i] - [i]')$            |
|----------|--------------------------------|---|
| <i>A</i> | <i>A</i> = 1.48                | <b>BD</b> + <i>CE</i> + <i>FG</i> = 19.15 |
| <i>B</i> | <b>B</b> = 38.05               | <i>AD</i> + <i>CF</i> + <i>EG</i> = 0.33  |
| <i>C</i> | <i>C</i> = -1.80               | <i>AE</i> + <i>BF</i> + <i>DG</i> = 1.53  |
| <i>D</i> | <b>D</b> = 29.38               | <i>AB</i> + <i>CG</i> + <i>EF</i> = -0.50 |
| <i>E</i> | <i>E</i> = 0.13                | <i>AC</i> + <i>BG</i> + <i>DF</i> = -0.40 |
| <i>F</i> | <i>F</i> = 0.50                | <i>BC</i> + <i>AG</i> + <i>DE</i> = -1.53 |
| <i>G</i> | <i>G</i> = 0.13                | <i>CD</i> + <i>BE</i> + <i>AF</i> = -2.55 |

The largest two effects are *B* and *D*. Furthermore, the third largest effect is *BD* + *CE* + *FG*, so it seems reasonable to attribute this to the *BD* interaction. The analyst used the two factors distance (*B*) and illumination level (*D*) in subsequent experiments with the other factors *A*, *C*, *E*, and *F* at standard settings and verified the results obtained here. He decided to use subjects as blocks in these new experiments rather than ignore a potential subject effect, because several different subjects had to be used to complete the experiment.

While the full fold-over strategy illustrated in Example 4.5 is used very frequently, there is another variation of this technique that occasionally proves helpful. In a **single-factor fold-over** we add to a fractional factorial design of resolution III a second fraction of the same size with the signs for only one of the factors reversed. In the combined design, we will be able to estimate the main effect of the factor for which the signs were reversed as well as all two-factor interactions involving that factor.

To illustrate, consider the  $2^{7-4}_{III}$  design in Table 4.15. Suppose that along with this principal fraction a second fractional design with the signs reversed in the column for factor *D* is also run. That is, the column for *D* in the second fraction is

$$- + + - - + + -$$

The effects that may be estimated from the first fraction are shown in Equation 4.1, and from the second fraction we obtain

$$\begin{aligned}
 [A]' &\longrightarrow A - BD + CE + FG \\
 [B]' &\longrightarrow B - AD + CF + EG \\
 [C]' &\longrightarrow C + AE + BF - DG \\
 [D]' &\longrightarrow D - AB - CG - EF
 \end{aligned}$$

that is,

$$\begin{aligned}
 [D]' &\longrightarrow -D + AB + CG + EF \\
 [E]' &\longrightarrow E + AC + BG - DF \\
 [F]' &\longrightarrow F + BC + AG - DE \\
 [G]' &\longrightarrow G - CD + BE + AF
 \end{aligned}$$

assuming that three-factor and higher interactions are insignificant. Now from the two linear combinations of effects  $\frac{1}{2}([i] + [i]')$  and  $\frac{1}{2}([i] - [i]')$  we obtain

| $i$ | $\frac{1}{2}([i] + [i]')$ | $\frac{1}{2}([i] - [i]')$ |
|-----|---------------------------|---------------------------|
| $A$ | $A + CE + FG$             | $BD$                      |
| $B$ | $B + CF + EG$             | $AD$                      |
| $C$ | $C + AE + BF$             | $DG$                      |
| $D$ | $D$                       | $AB + CG + EF$            |
| $E$ | $E + AC + BG$             | $DF$                      |
| $F$ | $F + BC + AG$             | $DE$                      |
| $G$ | $G + BE + AF$             | $CD$                      |

Thus, we have isolated the main effect of  $D$  and all of its two-factor interactions.

**The Defining Relation for a Fold-over Design** Combining fractional factorial designs via fold-over as demonstrated in Example 4.5 is a very useful technique. It is often of interest to know the defining relation for the combined design. Fortunately, this can be easily determined. Each separate fraction will have  $L + U$  words used as generators:  $L$  words of like sign and  $U$  words of unlike sign. The combined design will have  $L + U - 1$  words used as generators. These will be the  $L$  words of like sign and the  $U - 1$  words consisting of independent even products of the words of unlike sign. (Even products are words taken two at a time, four at a time, and so forth.)

To illustrate this procedure, consider the design in Example 4.5. For the first fraction, the generators are

$$I = ABD, \quad I = ACE, \quad I = BCF, \quad \text{and} \quad I = ABCG$$

and for the second fraction, they are

$$I = -ABD, \quad I = -ACE, \quad I = -BCF, \quad \text{and} \quad I = ABCG$$

Notice that in the second fraction we have switched the signs on the generators with an odd number of letters. Also, notice that  $L + U = 1 + 3 = 4$ . The combined design will have  $I = ABCG$  (the like-sign word) as a generator, along with two words that are independent even products of the words of unlike sign. For example, take  $I = ABD$  and  $I = ACE$ ; then  $I = (ABD)(ACE) = BCDE$  is a generator of the combined design. Also, take  $I = ABD$  and  $I = BCF$ ; then  $I = (ABD)(BCF) = ACDF$  is a generator of the combined design. The complete defining relation for the combined design is

$$I = ABCG = BCDE = ACDF = ADEG = BDFG = ABEF = CEFG$$

**TABLE 4.19 Plus and Minus Signs for the Plackett–Burman Designs**

---

|                     |   |                     |
|---------------------|---|---------------------|
| $k = 11, N = 12$    | + + - + + + - - - + -   |                     |
| $k = 19, N = 20$    | + + - - + + + + - + - + - - - - + + -                           |                     |
| $k = 23, N = 24$    | + + + + + + - + - + + - - + + - - + - - - -                     |                     |
| $k = 35, N = 36$    | - + - + + + + - - - + + + + + - + + - - - - + - + - + + - - + - |                     |
|                     | $k = 27, N = 28$  |                     |
| + - + + + + - - - - | - + - - - + - - - +   | + + - + - + + + - + |
| + + - + + + - - - - | - - + + - - - + - -   | - + + + + - + + -   |
| - + + + + + - - - - | + - - - + - - - + -   | + - + - + + - + +   |
| - - - + - + + + +   | - - + - + - - - -   | + - + + + - + - +   |
| - - - + + - + + + + | + - - - - + + - -   | + + - - + + + + -   |
| - - - - + + + + +   | - + - + - - - + -   | - + + + - + - + +   |
| + + + - - - + - +   | - - + - - - + - + -   | + - + + - + + + -   |
| + + + - - - + + -   | + - - + - - - - +   | + + - + + - - + +   |
| + + + - - - - + +   | - + - - - + - + -   | - + + - + + + - +   |

---

Because the defining relation for the combined design contains only four-letter words, the combined design is of resolution IV.

**Plackett–Burman Designs** These designs, developed by Plackett and Burman (1946), are two-level fractional designs for studying up to  $k = N - 1$  variables in  $N$  runs, where  $N$  is a multiple of 4. If  $N$  is a power of 2, these designs are identical to those presented earlier in this section. However, for  $N = 12, 20, 24, 28,$  and  $36,$  the Plackett–Burman designs are sometimes of interest.

The upper half of Table 4.19 presents rows of plus and minus signs that are used to construct the Plackett–Burman designs for  $N = 12, 20, 24,$  and  $36,$  whereas the lower half of the table presents blocks of plus and minus signs for constructing the design for  $N = 28.$  The designs for  $N = 12, 20, 24,$  and  $36$  are obtained by writing the appropriate row in Table 4.19 as a column (or row). A second column (or row) is then generated from this first one by moving the elements of the column (or row) down (or to the right) one position and placing the last element in the first position. A third column (or row) is produced from the second similarly, and the process continued until column (or row)  $k$  is generated. A row of minus signs is then added, completing the design. For  $N = 28,$  the three blocks  $X, Y,$  and  $Z$  are written down in the order

$$\begin{matrix} X & Y & Z \\ Z & X & Y \\ Y & Z & X \end{matrix}$$

and a row of minus signs is added to these 27 rows. The design for  $N = 12$  runs and  $k = 11$  factors is shown in Table 4.20, where the generator in Table 4.19 was used as a column.

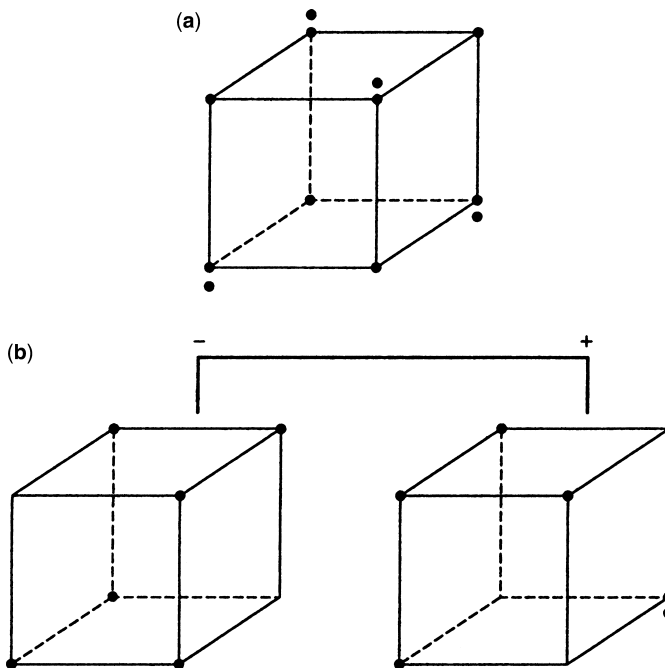
The Plackett–Burman designs for  $N = 12, 20, 24, 28,$  and  $36$  have somewhat complex alias structures. For example, in the 12-run design every main effect is **aliased** with every two-factor interaction not involving itself. For example, the main effect of  $A$  is aliased with the interactions  $BC, BD, BE,$  and so on. These are 45 two-factor interactions in this alias chain (and in every main effect alias chain). However, the constants in the alias chains are only  $\pm 1/3,$  not  $\pm 1.$  That is,  $A = \pm \frac{1}{3}BC \pm \frac{1}{3}BD \pm \dots$  This is often called **partial aliasing.** Plackett–Burman designs are an example of a **nonregular design**—that is, a design where

**TABLE 4.20 Plackett–Burman Design for  $N = 12, k = 11$  Using Generator in Table 4.19**

| Run | A | B | C | D | E | F | G | H | I | J | K |
|-----|---|---|---|---|---|---|---|---|---|---|---|
| 1   | + | - | + | - | - | - | + | + | + | - | + |
| 2   | + | + | - | + | - | - | - | + | + | + | - |
| 3   | - | + | + | - | + | - | - | - | + | + | + |
| 4   | + | - | + | + | - | + | - | - | - | + | + |
| 5   | + | + | - | + | + | - | + | - | - | - | + |
| 6   | + | + | + | - | + | + | - | + | - | - | - |
| 7   | - | + | + | + | - | + | + | - | + | - | - |
| 8   | - | - | + | + | + | - | + | + | - | + | - |
| 9   | - | - | - | + | + | + | - | + | + | - | + |
| 10  | + | - | - | - | + | + | + | - | + | + | - |
| 11  | - | + | - | - | - | + | + | + | - | + | + |
| 12  | - | - | - | - | - | - | - | - | - | - | - |

some of the constants in the alias relationships are not  $\pm 1$  or 0. In the larger designs, the situation is even more complex. As we will see in subsequent chapters, Plackett–Burman designs can be used in the construction of certain types of second-order response surface designs.

The projection properties of the Plackett–Burman designs are sometimes useful. For example, consider the 12-run design in Table 4.20. This design will project into three replicates of a full  $2^2$  design in any two of the original 11 factors. However, in three factors, the projected design is a full  $2^3$  factorial plus a  $2^{3-1}_{III}$  fractional factorial [see Fig. 4.17a].



**Figure 4.17** Projection of the 12-run Plackett–Burman design into three- and four-factor designs. (a) Projection into three factors. (b) Projection into four factors.

The four-dimensional projections are shown in Fig. 4.17b. Notice that these three- and four-factor projections are not balanced designs. This would potentially complicate their analysis. Plackett–Burman designs with  $N = 12, 20, 24, 28,$  and  $36$  are often called **nongeometric** Plackett–Burman designs.

### 4.6 RESOLUTION IV AND V DESIGNS

A  $2^{p-k}$  fractional factorial design is of resolution IV if the main effects are clear of two-factor interactions, but some two-factor interactions are aliased with each other. Thus, if three-factor and higher interactions are assumed negligible and suppressed, the main effects may be estimated directly in a  $2_{IV}^{k-p}$  design. An example is the  $2_{IV}^{6-2}$  design in Table 4.11. Furthermore, the two combined fractions of the  $2_{III}^{7-4}$  design in Example 4.5 yield a  $2_{IV}^{7-3}$  design.

Any  $2_{IV}^{k-p}$  design must contain at least  $2k$  runs. Resolution IV designs that contain exactly  $2k$  runs are called **minimal designs**. Resolution IV designs may be obtained from resolution III designs by the process of fold-over. Recall that to fold over a design, one simply adds to the original fraction a second fraction with all the signs reversed. Then the plus signs in the identity column  $I$  in the first fraction can be switched in the second fraction, and a  $(k + 1)$ st factor associated with this column. The result is a  $2_{IV}^{k+1-p}$  fractional factorial design. The process is demonstrated in Table 4.21 for the  $2_{III}^{3-1}$  design. It is easy to verify that the resulting design is a  $2_{IV}^{4-1}$  design with defining relation  $I = ABCD$ .

It is also possible to fold over resolution IV designs to separate two-factor interactions that are aliased with each other. One way to fold over a resolution IV design is to run a second fraction in which the sign is reversed on every design generator that has an *even number of letters*. To illustrate, consider the  $2_{IV}^{6-2}$  design used for the injection molding experiment in Example 4.3. The generators for the design in Table 4.13 are  $I = ABCE$  and  $I = BCDF$ . The second fraction uses the generators  $I = -ABCE$  and  $I = -BCDF$ , and the single generator for the combined design is  $I = ADEF$ . Thus, the combined design is still a resolution IV fractional factorial design. However, the alias relationships will be much simpler than in the original  $2_{IV}^{6-2}$  fractional. In fact, the only two-factor

**TABLE 4.21** A  $2_{IV}^{4-1}$  Design Obtained by Fold-Over

| $D$  | $A$ | $B$ | $C$ |
|--|-----|-----|-----|
| $I$  |     |     |     |
| Original $2_{III}^{3-1}$ with $I = ABC$    |     |     |     |
| +  | -   | -   | +   |
| +  | +   | -   | -   |
| +  | -   | +   | -   |
| +  | +   | +   | +   |
| Second $2_{III}^{3-1}$ with Signs Switched |     |     |     |
| -  | +   | +   | -   |
| -  | -   | +   | +   |
| -  | +   | -   | +   |
| -  | -   | -   | -   |

interactions that will be aliased are  $AD = EF$ ,  $AE = DF$ , and  $AF = DE$ . All the other two-factor interactions can be estimated from the combined design. The approach used often in practice is to construct the fold-over design by switching the signs on a single column in the original design. This causes all of the alias chains involving that factor to be broken.

Notice that when you start with a resolution III design, the fold-over procedure guarantees that the combined design will be of resolution IV, thereby ensuring that all the main effects can be separated from their two-factor interaction aliases. When folding over a resolution IV design, we will not necessarily separate all the two-factor interactions. In fact, if the original fraction has an alias structure with more than two two-factor interactions in any alias chain, folding over will not completely separate all the two-factor interactions. Notice that in the foregoing example, the  $2_{IV}^{6-2}$  has one such two-factor interaction alias chain. For more information on fold-over of resolution IV designs, see Montgomery and Runger (1996) and Montgomery (2013).

Resolution V designs are fractional factorials in which the main effects and the two-factor interactions do not have other main effects and two-factor interactions as their aliases. These are very powerful designs, allowing the unique estimation of all the main effects and two-factor interactions provided that all the three-factor and higher interactions are negligible. The smallest word in the defining relation of such a design must have five letters. The  $2^{5-1}$  with the generating relation  $I = ABCDE$  is of resolution V. Another example is the  $2_{V}^{8-2}$  design with the generating relations  $I = ABCDG$  and  $I = ABEFH$ . Further examples of these designs are given by Box and Hunter (1961b).

#### 4.7 ALIAS STRUCTURES IN FRACTIONAL FACTORIAL AND OTHER DESIGNS

We have shown how to find the alias relationships in a  $2^{k-p}$  fractional factorial design by use of the complete defining relation. This method works well in simple designs, such as the regular fractions we use most frequently, but it does not work as well in more complex settings, such as some of the nonregular fractions and nonorthogonal designs that we will discuss subsequently. Furthermore, there are some fractional factorials that do not have defining relations, so the defining relation method will not work for these types of designs at all.

Fortunately, there is a general method available that works satisfactorily in many situations. The method uses the polynomial or regression model representation of the model, say

$$\mathbf{y} = \mathbf{X}_1 \boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}$$

where  $\mathbf{X}_1$  is the design matrix for the experiment that has been conducted expanded to model form,  $\boldsymbol{\beta}_1$  is the vector of model parameters, and  $\boldsymbol{\varepsilon}$  is the usual vector of  $\text{NID}(0, \sigma^2)$  errors. Suppose that the true model is

$$\mathbf{y} = \mathbf{X}_1 \boldsymbol{\beta}_1 + \mathbf{X}_2 \boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$$



where the columns of  $\mathbf{X}_2$  contain additional factors not included in the original model (such as interactions) and  $\beta_2$  is the corresponding vector of model parameters. We can show that the expected value of  $\hat{\beta}_1$ , the least squares estimate of  $\beta_1$ , is

$$E(\hat{\beta}_1) = \beta_1 + (\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1\mathbf{X}_2\beta_2 = \beta_1 + \mathbf{A}\beta_2$$

The **alias matrix**  $\mathbf{A} = (\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1\mathbf{X}_2$  shows how estimates of terms in the fitted model are biased by active terms that are not in the fitted model. Each row of  $\mathbf{A}$  is associated with a parameter in the fitted model. Nonzero elements in a row of  $\mathbf{A}$  show the degree of biasing of the fitted model parameter due to terms associated with the columns of  $\mathbf{X}_2$ .

We illustrate the application of this procedure with a familiar example. Suppose that we have conducted a  $2^{3-1}$  design with defining relation  $I = ABC$  or  $I = x_1x_2x_3$ . The model that the experimenter plans to fit is the main-effects-only model

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \varepsilon$$

In the notation used above,

$$\beta_1 = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \quad \text{and} \quad \mathbf{X}_1 = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Suppose that the true model contains all the two-factor interaction so that

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{12}x_1x_2 + \beta_{13}x_1x_3 + \beta_{23}x_2x_3 + \varepsilon$$

and

$$\beta_2 = \begin{bmatrix} \beta_{12} \\ \beta_{13} \\ \beta_{23} \end{bmatrix} \quad \text{and} \quad \mathbf{X}_2 = \begin{bmatrix} 1 & -1 & -1 \\ -1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$$

Now

$$\mathbf{X}'_1\mathbf{X}_1 = 4\mathbf{I}_4 \quad \text{and} \quad \mathbf{X}'_1\mathbf{X}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \\ 4 & 0 & 0 \end{bmatrix}$$

Therefore,

$$(\mathbf{X}'_1\mathbf{X}_1)^{-1} = \frac{1}{4}\mathbf{I}_4$$

and

$$\begin{aligned}
 E(\beta_1) &= \beta_1 + A\beta_2 \\
 E \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} &= \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \frac{1}{4}I_4 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \\ 4 & 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_{12} \\ \beta_{13} \\ \beta_{23} \end{bmatrix} \\
 &= \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_{12} \\ \beta_{13} \\ \beta_{23} \end{bmatrix} \\
 &= \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} 0 \\ \beta_{23} \\ \beta_{13} \\ \beta_{12} \end{bmatrix} \\
 &= \begin{bmatrix} \beta_0 \\ \beta_1 + \beta_{23} \\ \beta_2 + \beta_{13} \\ \beta_3 + \beta_{12} \end{bmatrix}
 \end{aligned}$$

The interpretation of this, of course, is that each of the main effects is aliased with one of the two-factor interpretations, which we know to be the case for this design. Notice that every row of the alias matrix represents one of the factors in  $\beta_1$  and that every column represents one of the factors in  $\beta_2$  while this is a very simple example, the method is very general and can be applied to much more complex designs.

JMP will display the alias matrix for fractional factorial designs. For the  $2^{3-1}$  design the JMP display of the alias matrix is

| Effect    | X1*X2 | X1*X3 | X2*X3 |
|-----------|-------|-------|-------|
| Intercept | 0     | 0     | 0     |
| X1        | 0     | 0     | 1     |
| X2        | 0     | 1     | 0     |
| X3        | 1     | 0     | 0     |

The 1 in the second row of the matrix indicates that the main effect of factor A (or X<sub>1</sub>) is aliased with the X<sub>2</sub>X<sub>3</sub> interaction.

#### 4.8 NONREGULAR FRACTIONAL FACTORIAL DESIGNS

The  $2^{k-p}$  fractional factorial designs discussed in previous sections are a staple for factor screening in modern industrial applications. In these designs if factors are aliased with other factors, then the aliasing is complete; that is, the non-zero constants in the alias relationships are either +1 or -1. Resolution IV designs are particularly popular because they avoid the confounding of main effects and two-factor interactions found in resolution III designs while avoiding the larger sample size requirements of resolution V designs. However, when the number of factors is relatively large, say  $k = 9$  or more, resolution III designs are widely used.

The two-factor interaction aliasing in the resolution III and IV designs can result in experiments whose outcomes have ambiguous conclusions. For example, suppose that in the injection molding experiment in Example 4.3 the main effects of the factors A, B, C, and E were found to be important along with one two-factor interaction alias chain  $AB + CE$ . Without external process knowledge, the experimenter could not decide whether the AB interaction, the CE interaction, or some linear combination of them represents the true state of nature. To resolve this ambiguity requires additional runs. Fold-over could be used to resolve this ambiguity. The fold-over design could be constructed by changing the signs in column A to produce an additional 16-run design.

Another strategy would be to consider a **partial fold-over**; this approach would result in an additional 8-run design. Montgomery (2013) provides the details of constructing partial fold overs along with an example. Yet another approach would be to augment the original design with additional runs selected by a  $D$ -optimal design construction algorithm, where the model of interest would contain both the AB and the CE interaction.

While strong two-factor interactions may be less likely than strong main effects, there are likely to be many more interactions than main effects in screening situations. As a result, the likelihood of at least one significant interaction effect is quite high. There is often substantial reluctance to commit additional time and material to a study with unclear results. Consequently, experimenters often want to avoid the need for a follow-up study. In this section we show how specific choices of **nonregular** two-level fractional factorial designs can be used in experiments with between 6 and 14 factors and potentially avoid subsequent experimentation when two-factor interactions are active. Section 4.5 introduced the class of Plackett–Burman designs and noted that they were nonregular designs because the constants in the alias relationships were not all  $\pm 1$ . Section 4.7.1 presents nonregular designs for 6, 7, and 8 factors in 16 runs. These designs have no complete confounding of pairs of two-factor interactions. These designs are excellent alternatives for the regular minimum aberration resolution IV fractional factorials. In Section 4.7.2 we present nonregular designs for between 9 and 14 factors in 16 runs that have no complete aliasing of main effects and two-factor interactions. These designs are alternative to the regular minimum aberration resolution III fractional factorials. We also present metrics to evaluate these fractional factorial designs, show how the recommended nonregular 16-run designs were obtained, and discuss analysis methods.

Screening designs are primarily concerned with the discovery of active factors. This factor activity generally expresses itself through a main effect or a factor's involvement in a two-factor interaction. Consider the model

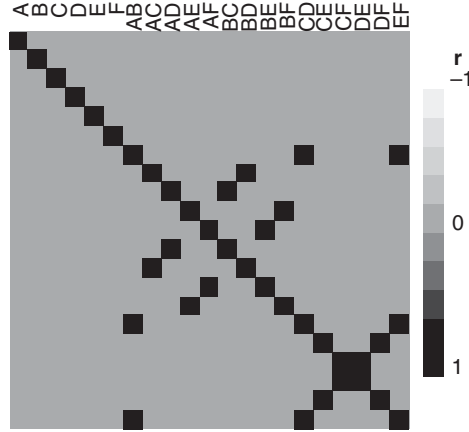
$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (4.2)$$

where  $\mathbf{X}$  contains columns for the intercept, main effects, and all two-factor interactions,  $\boldsymbol{\beta}$  is the vector of model parameters, and  $\boldsymbol{\varepsilon}$  is the usual vector of  $\text{NID}(0, \sigma^2)$  random errors.

Consider the case of six factors in 16 runs and a model with all main effects and two-factor interactions. For this situation the  $\mathbf{X}$  matrix has more columns than rows. Thus, it is not of full rank and the usual least squares estimate for  $\boldsymbol{\beta}$  does not exist because the matrix  $\mathbf{X}'\mathbf{X}$  is singular. With respect to this model, every 16-run design is **supersaturated**. Booth and Cox (1962) introduced the  $E(s^2)$  criterion as a diagnostic measure for comparing supersaturated designs, where

$$E(s^2) = \sum_{i < j}^2 (x_i' x_j) / (k(k - 1))$$

and  $k$  is the number of columns in  $\mathbf{X}$ .



**Figure 4.18** The correlation matrix for the regular  $2^{6-2}$  resolution IV fractional factorial design.

Minimizing the  $E(s^2)$  criterion is equivalent to minimizing the sum of squared off-diagonal elements of the correlation matrix of  $\mathbf{X}$ . Removing the constant column from  $\mathbf{X}$ , the correlation matrix of the regular  $2^{6-2}$  resolution IV 16-run six-factor design is  $21 \times 21$  with one row and column for each of the six main effects and 15 two-factor interactions. Figure 4.18 shows the **cell plot** of the **correlation matrix** for the principal fraction of this design. In Fig. 4.18 we note that the correlation is zero between all main effects and two-factor interactions (because the design is resolution IV) and that the correlation is +1 between every two-factor interaction and at least one other two-factor interaction. These two-factor interactions are completely confounded. If another member of the same design family had been used, at least one of the generators would have been used with a negative sign in design construction and some of the entries of the correlation matrix would have been  $-1$ . There still would be complete confounding of two-factor interactions in the design.

Jones and Montgomery (2010) introduced the cell plot of the correlation matrix as a useful graphical way to show the aliasing relationships in fractional factorials and to compare nonregular designs to their regular fractional factorial counterparts. In Fig. 4.18 it is a display of the confounding pattern, much like can be seen in the alias matrix. We introduced the alias matrix earlier in this chapter. Recall that we plan to fit the model

$$y = \mathbf{X}_1 \boldsymbol{\beta}_1 + \varepsilon$$

where  $\mathbf{X}_1$  is the design matrix for the experiment that has been conducted expanded to model form,  $\boldsymbol{\beta}_1$  is the vector of model parameters, and  $\varepsilon$  is the usual vector of  $NID(0, \sigma^2)$  errors but that the true model is

$$y = \mathbf{X}_1 \boldsymbol{\beta}_1 + \mathbf{X}_2 \boldsymbol{\beta}_2 + \varepsilon$$

where the columns of  $\mathbf{X}_2$  contain additional factors not included in the original model (such as interactions) and  $\boldsymbol{\beta}_2$  is the corresponding vector of model parameters. Earlier we observed that the expected value of  $\hat{\boldsymbol{\beta}}_1$ , the least squares estimate of  $\boldsymbol{\beta}_1$ , is

$$E(\hat{\boldsymbol{\beta}}_1) = \boldsymbol{\beta}_1 + (\mathbf{X}'_1 \mathbf{X}_1)^{-1} \mathbf{X}'_1 \mathbf{X}_2 \boldsymbol{\beta}_2 = \boldsymbol{\beta}_1 + \mathbf{A} \boldsymbol{\beta}_2$$



**TABLE 4.23 The Hall II Design**

| Run | A  | B  | C  | D  | E  | F  | G  | H  | J  | K  | L  | M  | N  | P  | Q  |
|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 2   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |
| 3   | 1  | 1  | 1  | -1 | -1 | -1 | -1 | 1  | 1  | 1  | 1  | -1 | -1 | -1 | -1 |
| 4   | 1  | 1  | 1  | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | 1  | 1  | 1  | 1  |
| 5   | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 |
| 6   | 1  | -1 | -1 | 1  | 1  | -1 | -1 | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  |
| 7   | 1  | -1 | -1 | -1 | -1 | 1  | 1  | 1  | 1  | -1 | -1 | -1 | -1 | 1  | 1  |
| 8   | 1  | -1 | -1 | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | 1  | 1  | -1 | -1 |
| 9   | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 |
| 10  | -1 | 1  | -1 | 1  | -1 | 1  | -1 | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  |
| 11  | -1 | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 | -1 | 1  | -1 | 1  |
| 12  | -1 | 1  | -1 | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 |
| 13  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 |
| 14  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 | -1 | 1  |
| 15  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  |
| 16  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 |

**TABLE 4.24 The Hall III Design**

| Run | A  | B  | C  | D  | E  | F  | G  | H  | J  | K  | L  | M  | N  | P  | Q  |
|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 2   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |
| 3   | 1  | 1  | 1  | -1 | -1 | -1 | -1 | 1  | 1  | 1  | 1  | -1 | -1 | -1 | -1 |
| 4   | 1  | 1  | 1  | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | 1  | 1  | 1  | 1  |
| 5   | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 |
| 6   | 1  | -1 | -1 | 1  | 1  | -1 | -1 | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  |
| 7   | 1  | -1 | -1 | -1 | -1 | 1  | 1  | 1  | 1  | -1 | -1 | -1 | -1 | 1  | 1  |
| 8   | 1  | -1 | -1 | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | 1  | 1  | -1 | -1 |
| 9   | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 |
| 10  | -1 | 1  | -1 | 1  | -1 | 1  | -1 | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  |
| 11  | -1 | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  |
| 12  | -1 | 1  | -1 | -1 | 1  | -1 | 1  | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 |
| 13  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 |
| 14  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 | -1 | 1  |
| 15  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 | 1  | -1 | -1 | 1  | -1 | 1  |
| 16  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 |

**TABLE 4.25 The Hall IV Design**

| Run | A  | B  | C  | D  | E  | F  | G  | H  | J  | K  | L  | M  | N  | P  | Q  |
|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 2   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |
| 3   | 1  | 1  | 1  | -1 | -1 | -1 | -1 | 1  | 1  | 1  | 1  | -1 | -1 | -1 | -1 |
| 4   | 1  | 1  | 1  | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | 1  | 1  | 1  | 1  |
| 5   | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 |
| 6   | 1  | -1 | -1 | 1  | 1  | -1 | -1 | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  |
| 7   | 1  | -1 | -1 | -1 | -1 | 1  | 1  | 1  | 1  | -1 | -1 | -1 | -1 | 1  | 1  |
| 8   | 1  | -1 | -1 | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | 1  | 1  | -1 | -1 |
| 9   | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 |
| 10  | -1 | 1  | -1 | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | -1 | 1  | -1 | 1  |
| 11  | -1 | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | -1 | 1  |
| 12  | -1 | 1  | -1 | -1 | 1  | -1 | 1  | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 |
| 13  | -1 | -1 | 1  | 1  | -1 | 1  | -1 | -1 | 1  | -1 | 1  | -1 | 1  | 1  | -1 |
| 14  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 | -1 | 1  |
| 15  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 | 1  | -1 | -1 | 1  | -1 | 1  |
| 16  | -1 | -1 | 1  | -1 | 1  | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | 1  | -1 |

**TABLE 4.26 The Hall V Design**

| Run | A  | B  | C  | D  | E  | F  | G  | H  | J  | K  | L  | M  | N  | P  | Q  |
|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 2   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |
| 3   | 1  | 1  | 1  | -1 | -1 | -1 | -1 | 1  | 1  | 1  | 1  | -1 | -1 | -1 | -1 |
| 4   | 1  | 1  | 1  | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | 1  | 1  | 1  | 1  |
| 5   | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 |
| 6   | 1  | -1 | -1 | 1  | 1  | -1 | -1 | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  |
| 7   | 1  | -1 | -1 | -1 | -1 | 1  | 1  | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 |
| 8   | 1  | -1 | -1 | -1 | -1 | 1  | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  |
| 9   | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  | 1  | -1 | -1 | -1 | -1 | 1  | 1  |
| 10  | -1 | 1  | -1 | 1  | -1 | 1  | -1 | -1 | -1 | 1  | 1  | 1  | 1  | -1 | -1 |
| 11  | -1 | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 |
| 12  | -1 | 1  | -1 | -1 | 1  | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 | -1 | 1  |
| 13  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | 1  | -1 | -1 | 1  | -1 | 1  |
| 14  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 |
| 15  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  |
| 16  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 |

**TABLE 4.27 Number of 16-Run Orthogonal Nonisomorphic Designs**

| Number of Factors | Number of Designs |
|-------------------|-------------------|
| 6                 | 27                |
| 7                 | 55                |
| 8                 | 80                |

**TABLE 4.28 A Nonregular Orthogonal Design for  $k = 6$  Factors in 16 Runs**

| Run | A  | B  | C  | D  | E  | F  |
|-----|----|----|----|----|----|----|
| 1   | 1  | 1  | 1  | 1  | 1  | 1  |
| 2   | 1  | 1  | -1 | -1 | -1 | -1 |
| 3   | -1 | -1 | 1  | 1  | -1 | -1 |
| 4   | -1 | -1 | -1 | -1 | 1  | 1  |
| 5   | 1  | 1  | 1  | -1 | 1  | -1 |
| 6   | 1  | 1  | -1 | 1  | -1 | 1  |
| 7   | -1 | -1 | 1  | -1 | -1 | 1  |
| 8   | -1 | -1 | -1 | 1  | 1  | -1 |
| 9   | 1  | -1 | 1  | 1  | 1  | -1 |
| 10  | 1  | -1 | -1 | -1 | -1 | 1  |
| 11  | -1 | 1  | 1  | 1  | -1 | 1  |
| 12  | -1 | 1  | -1 | -1 | 1  | -1 |
| 13  | 1  | -1 | 1  | -1 | -1 | -1 |
| 14  | 1  | -1 | -1 | 1  | 1  | 1  |
| 15  | -1 | 1  | 1  | -1 | 1  | 1  |
| 16  | -1 | 1  | -1 | 1  | -1 | -1 |

The nonregular designs that Jones and Montgomery recommended for 6, 7, and 8 factors are shown in Tables 4.28, 4.29, and 4.30. The six-factor design in Table 4.28 is found from columns D, E, H, K, M and Q of Hall II. The correlation matrix for this design along with the correlation matrix for the corresponding regular fraction is in Figure 4.19. Notice that like the regular  $2^{6-2}$  the design in Table 4.28 is first-order orthogonal but unlike the

**TABLE 4.29 A Nonregular Orthogonal Design for  $k = 7$  Factors in 16 Runs**

| Run | A  | B  | C  | D  | E  | F  | G  |
|-----|----|----|----|----|----|----|----|
| 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 2   | 1  | 1  | 1  | -1 | -1 | -1 | -1 |
| 3   | 1  | 1  | -1 | 1  | 1  | -1 | -1 |
| 4   | 1  | 1  | -1 | -1 | -1 | 1  | 1  |
| 5   | 1  | -1 | 1  | 1  | -1 | 1  | -1 |
| 6   | 1  | -1 | 1  | -1 | 1  | -1 | 1  |
| 7   | 1  | -1 | -1 | 1  | -1 | -1 | 1  |
| 8   | 1  | -1 | -1 | -1 | 1  | 1  | -1 |
| 9   | -1 | 1  | 1  | 1  | 1  | 1  | -1 |
| 10  | -1 | 1  | 1  | -1 | -1 | -1 | 1  |
| 11  | -1 | 1  | -1 | 1  | -1 | 1  | 1  |
| 12  | -1 | 1  | -1 | -1 | 1  | -1 | -1 |
| 13  | -1 | -1 | 1  | 1  | -1 | -1 | -1 |
| 14  | -1 | -1 | 1  | -1 | 1  | 1  | 1  |
| 15  | -1 | -1 | -1 | 1  | 1  | -1 | 1  |
| 16  | -1 | -1 | -1 | -1 | -1 | 1  | -1 |



TABLE 4.30 A Nonregular Orthogonal Design for  $k = 8$  Factors in 16 Runs

| Run | A  | B  | C  | D  | E  | F  | G  | H  |
|-----|----|----|----|----|----|----|----|----|
| 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 2   | 1  | 1  | 1  | 1  | -1 | -1 | -1 | -1 |
| 3   | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 |
| 4   | 1  | 1  | -1 | -1 | -1 | -1 | 1  | 1  |
| 5   | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 |
| 6   | 1  | -1 | 1  | -1 | -1 | 1  | -1 | 1  |
| 7   | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  |
| 8   | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 |
| 9   | -1 | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 10  | -1 | 1  | 1  | -1 | 1  | -1 | -1 | -1 |
| 11  | -1 | 1  | -1 | 1  | -1 | -1 | 1  | -1 |
| 12  | -1 | 1  | -1 | -1 | -1 | 1  | -1 | 1  |
| 13  | -1 | -1 | 1  | 1  | -1 | -1 | -1 | 1  |
| 14  | -1 | -1 | 1  | -1 | -1 | 1  | 1  | -1 |
| 15  | -1 | -1 | -1 | 1  | 1  | 1  | -1 | -1 |
| 16  | -1 | -1 | -1 | -1 | 1  | -1 | 1  | 1  |

regular design, there are no two-factor interactions that are aliased with each other. All of the off-diagonal entries in the correlation matrix are either zero,  $-0.5$ , or  $+0.5$ . Because there is no complete confounding of two-factor interactions, Jones and Montgomery called this nonregular fraction a **no-confounding design**.

Table 4.29 presents the recommended seven-factor 16-run design. This design was constructed by selecting columns A, B, D, H, J, M, and Q from Hall III. The correlation matrix for this design and the regular  $2^{7-3}$  fraction is shown in Figure 4.20. The no-confounding design is first-order orthogonal and there is no complete confounding of

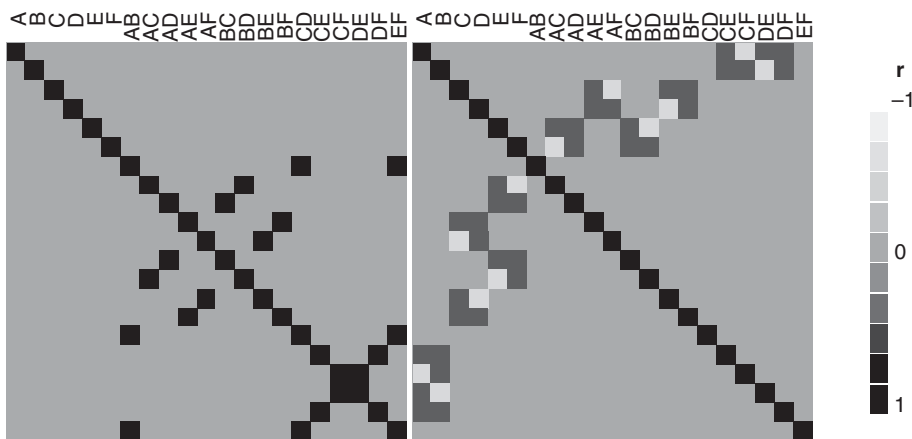
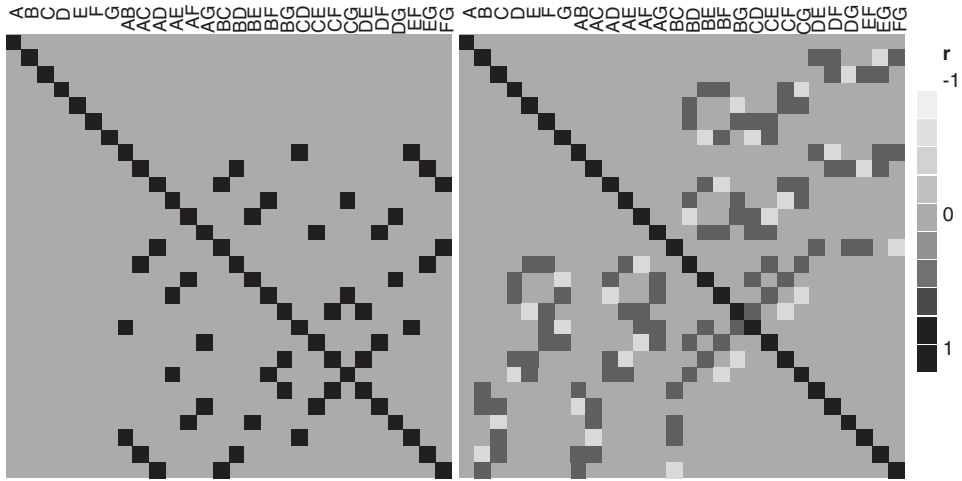


Figure 4.19 Correlation matrix (a) regular  $2^{6-2}$  fractional factorial. (b) Nonregular no-confounding design.

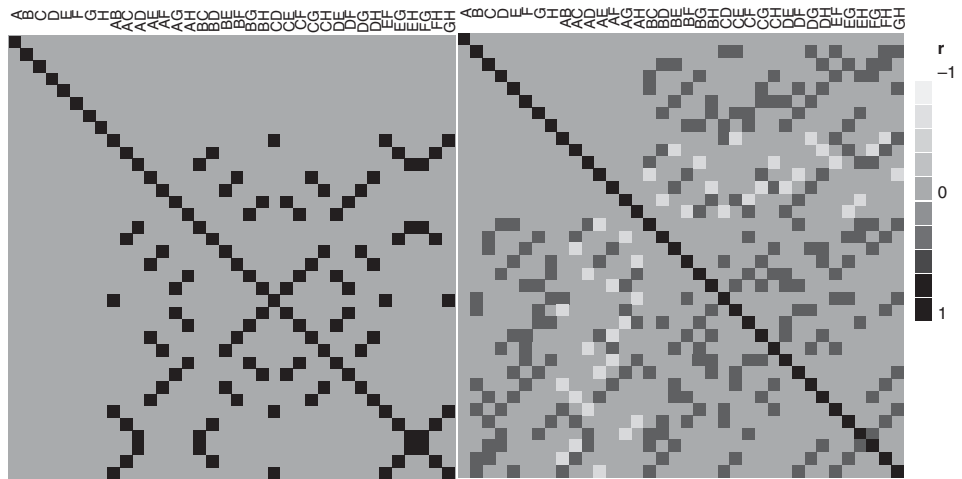


**Figure 4.20** Correlation matrix (a) regular  $2^{7-3}$  fractional factorial. (b) Nonregular no-confounding design.

two-factor interactions. All off-diagonal elements of the correlation matrix are either zero,  $-0.5$ , or  $+0.5$ .

Table 4.30 presents the recommended eight-factor 16-run design. This design was constructed by choosing columns A, B, D, F, H, J, M, and O from Hall IV. The correlation matrix for this design and the regular  $2^{8-4}$  fraction is shown in Fig. 4.21. The no-confounding design is orthogonal for the first-order model and there is no complete confounding of two-factor interactions. All off-diagonal elements of the correlation matrix are either zero,  $-0.5$ , or  $+0.5$ .

Table 4.31 compares the popular minimum aberration resolution IV designs to the nonregular alternatives designs on the metrics described previously. As shown in the cell plots



**Figure 4.21** Correlation matrix (a) Regular  $2^{8-4}$  fractional factorial. (b) Nonregular no-confounding design.

**TABLE 4.31 Design Comparison on Metrics**

| $N$ Factors | Design        | Confounded Effect Pairs | $E(s^2)$ | Trace( $\mathbf{A}\mathbf{A}'$ ) |
|-------------|---------------|-------------------------|----------|----------------------------------|
| 6           | Recommended   | 0                       | 7.31     | 6                                |
|             | Resolution IV | 9                       | 10.97    | 0                                |
| 7           | Recommended   | 0                       | 10.16    | 6                                |
|             | Resolution IV | 21                      | 14.20    | 0                                |
| 8           | Recommended   | 0                       | 12.80    | 10.5                             |
|             | Resolution IV | 42                      | 17.07    | 0                                |

of the correlation matrices, the recommended designs outperform the minimum aberration designs for the number of confounded pairs of effects. They also are substantially better with respect to the  $E(s^2)$  criterion. The recommended designs all achieve the minimum value of the trace criterion for all of the possible nonregular designs. The price that the Jones and Montgomery recommended designs pay for avoiding any pure confounding is that there is some correlation between main effects and two-factor interactions.

Jones and Nachtsheim (2011a,b) have proposed an interesting and useful design optimality criterion intended to trade off variance and bias in a design. Their criterion is to choose a design to minimize

$$\begin{aligned} & \text{trace}(\mathbf{A}'\mathbf{A}) \\ & \text{subject to} \\ & D_e \geq l_D \end{aligned} \tag{4.3}$$

where  $D_e$  is the  $D$ -efficiency of the design and  $l_D$  is a lower bound on the  $D$ -efficiency. The bias component of the trade-off is reflected in the trace objective function, and the variance component of the trade-off is reflected in the lower bound on the  $D$ -efficiency. Jones and Nachtsheim give an efficient algorithm for implementing this optimality criterion. The nonconfounding designs presented above can all be constructed using this algorithm. It turns out that all of the designs in Tables 4.28, 4.29, and 4.30 achieve the minimum value of the trace and have a  $D$ -efficiency of 100 percent; that is, they are first-order orthogonal designs. Johnson and Jones (2010) show how these designs can be constructed using a version of the column generator technique.

The designs in Tables 4.28, 4.29, and 4.30 have useful projection properties. They either project into a 16-run full factorial in groups of 4 factors or fractions with at least 12 distinct design points. Therefore, they can fit all main effect plus two-factor interaction models with up to  $k = 4$  factors. For more information and examples, see Shinde, Montgomery, and Jones (2014) and Montgomery (2013).

#### 4.8.2 Nonregular Fractional Factorial Designs for 9 Through 14 Factors in 16 Runs

Resolution III fractional factorial designs are popular for factor screening problems in situations where there are a moderate-to-large number of factors because these designs contain a relatively small number of runs and they are effective in identifying the unimportant factors and elevating potentially important factors for further experimentation. Designs in 16 runs are extremely popular because the number of runs is usually within the resources available to most experimenters.

**TABLE 4.32 Number of Nonisomorphic Nonregular 16-Run Designs**

| Number of Factors | Number of Nonisomorphic Designs |
|-------------------|---------------------------------|
| 6                 | 27                              |
| 7                 | 55                              |
| 8                 | 80                              |
| 9                 | 87                              |
| 10                | 78                              |
| 11                | 58                              |
| 12                | 36                              |
| 13                | 18                              |
| 14                | 10                              |
| 15                | 5                               |

Because the regular resolution III designs alias main effects and two-factor interactions, and the aliased effects are completely confounded, experimenters often end up with ambiguous conclusions about which main effects and two-factor interactions are important. Resolving these ambiguities requires either additional experimentation (such as use of a fold-over design to augment the original fraction) or assumptions about which effects are important or external process knowledge. This is very similar to the situation encountered in the previous section, except now main effects are completely confounded with two-factor interactions. Just as in that section, it is possible to develop no-confounding designs for 9–14 factors in 16 runs that are good alternatives to the usual minimum aberration resolution III designs when there are only a few main effects and two-factor interactions that are important. Table 4.32 is an extension of Table 4.27, showing all possible nonisomorphic nonregular 16-run designs with 6 to 15 factors. The recommended designs in Tables 4.33 through 4.38 are chosen from the designs in this table. They are projections

**TABLE 4.33 Recommended 16-Run 9-Factor No-Confounding Design**

| Run | A  | B  | C  | D  | E  | F  | G  | H  | J  |
|-----|----|----|----|----|----|----|----|----|----|
| 1   | -1 | -1 | -1 | -1 | -1 | -1 | 1  | -1 | 1  |
| 2   | -1 | -1 | -1 | 1  | -1 | 1  | -1 | 1  | -1 |
| 3   | -1 | -1 | 1  | -1 | 1  | 1  | 1  | 1  | -1 |
| 4   | -1 | -1 | 1  | 1  | 1  | -1 | -1 | -1 | 1  |
| 5   | -1 | 1  | -1 | -1 | 1  | 1  | -1 | 1  | 1  |
| 6   | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 | -1 |
| 7   | -1 | 1  | 1  | -1 | -1 | -1 | -1 | 1  | -1 |
| 8   | -1 | 1  | 1  | 1  | -1 | 1  | 1  | -1 | 1  |
| 9   | 1  | -1 | -1 | -1 | 1  | -1 | -1 | -1 | -1 |
| 10  | 1  | -1 | -1 | 1  | 1  | 1  | 1  | 1  | 1  |
| 11  | 1  | -1 | 1  | -1 | -1 | 1  | -1 | -1 | 1  |
| 12  | 1  | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 |
| 13  | 1  | 1  | -1 | -1 | -1 | 1  | 1  | -1 | -1 |
| 14  | 1  | 1  | -1 | 1  | -1 | -1 | -1 | 1  | 1  |
| 15  | 1  | 1  | 1  | -1 | 1  | -1 | 1  | 1  | 1  |
| 16  | 1  | 1  | 1  | 1  | 1  | 1  | -1 | -1 | -1 |

**TABLE 4.34 Recommended 16-Run 10-Factor No-Confounding Design**

| Run | A  | B  | C  | D  | E  | F  | G  | H  | J  | K  |
|-----|----|----|----|----|----|----|----|----|----|----|
| 1   | -1 | -1 | -1 | -1 | 1  | -1 | -1 | 1  | -1 | 1  |
| 2   | -1 | -1 | -1 | 1  | 1  | 1  | -1 | -1 | 1  | 1  |
| 3   | -1 | -1 | 1  | -1 | -1 | 1  | 1  | 1  | 1  | 1  |
| 4   | -1 | -1 | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 |
| 5   | -1 | 1  | -1 | 1  | -1 | 1  | 1  | 1  | -1 | 1  |
| 6   | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 | -1 | -1 |
| 7   | -1 | 1  | 1  | -1 | -1 | -1 | -1 | 1  | -1 | -1 |
| 8   | -1 | 1  | 1  | 1  | -1 | 1  | -1 | -1 | 1  | -1 |
| 9   | 1  | -1 | -1 | -1 | -1 | 1  | 1  | -1 | -1 | -1 |
| 10  | 1  | -1 | -1 | 1  | -1 | -1 | -1 | 1  | 1  | -1 |
| 11  | 1  | -1 | 1  | 1  | -1 | -1 | 1  | -1 | -1 | 1  |
| 12  | 1  | -1 | 1  | 1  | 1  | 1  | -1 | 1  | -1 | -1 |
| 13  | 1  | 1  | -1 | -1 | -1 | -1 | -1 | -1 | 1  | 1  |
| 14  | 1  | 1  | -1 | -1 | 1  | 1  | 1  | 1  | 1  | -1 |
| 15  | 1  | 1  | 1  | -1 | 1  | 1  | -1 | -1 | -1 | 1  |
| 16  | 1  | 1  | 1  | 1  | 1  | -1 | 1  | 1  | 1  | 1  |

**TABLE 4.35 Recommended 16-Run 11-Factor No-Confounding Design**

| Run | A  | B  | C  | D  | E  | F  | G  | H  | J  | K  | L  |
|-----|----|----|----|----|----|----|----|----|----|----|----|
| 1   | -1 | -1 | -1 | 1  | 1  | -1 | -1 | -1 | 1  | -1 | 1  |
| 2   | -1 | -1 | 1  | -1 | -1 | -1 | 1  | -1 | 1  | -1 | -1 |
| 3   | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 | -1 | -1 |
| 4   | -1 | -1 | 1  | 1  | -1 | 1  | 1  | 1  | -1 | 1  | 1  |
| 5   | -1 | 1  | -1 | -1 | -1 | -1 | -1 | -1 | -1 | 1  | 1  |
| 6   | -1 | 1  | -1 | 1  | -1 | 1  | 1  | -1 | -1 | -1 | -1 |
| 7   | -1 | 1  | -1 | 1  | 1  | 1  | -1 | 1  | 1  | 1  | -1 |
| 8   | -1 | 1  | 1  | -1 | 1  | -1 | 1  | 1  | 1  | 1  | 1  |
| 9   | 1  | -1 | -1 | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  |
| 10  | 1  | -1 | -1 | -1 | 1  | 1  | 1  | -1 | -1 | 1  | 1  |
| 11  | 1  | -1 | -1 | 1  | -1 | -1 | 1  | 1  | 1  | 1  | -1 |
| 12  | 1  | -1 | 1  | 1  | 1  | -1 | -1 | -1 | -1 | 1  | -1 |
| 13  | 1  | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | -1 | -1 |
| 14  | 1  | 1  | 1  | -1 | -1 | 1  | -1 | -1 | 1  | 1  | -1 |
| 15  | 1  | 1  | 1  | 1  | -1 | -1 | -1 | 1  | -1 | -1 | 1  |
| 16  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | -1 | 1  | -1 | 1  |

**TABLE 4.36 Recommended 16-Run 12-Factor No-Confounding Design**

| Run | A  | B  | C  | D  | E  | F  | G  | H  | J  | K  | L  | M  |
|-----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1   | -1 | -1 | -1 | -1 | 1  | -1 | -1 | 1  | 1  | -1 | 1  | 1  |
| 2   | -1 | -1 | -1 | 1  | -1 | 1  | 1  | 1  | -1 | -1 | 1  | -1 |
| 3   | -1 | -1 | 1  | -1 | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  |
| 4   | -1 | -1 | 1  | 1  | 1  | 1  | -1 | -1 | -1 | 1  | -1 | -1 |
| 5   | -1 | 1  | -1 | 1  | -1 | -1 | -1 | -1 | -1 | -1 | -1 | 1  |
| 6   | -1 | 1  | -1 | 1  | 1  | 1  | 1  | -1 | 1  | 1  | 1  | 1  |
| 7   | -1 | 1  | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | -1 | -1 |
| 8   | -1 | 1  | 1  | -1 | 1  | -1 | 1  | 1  | -1 | 1  | 1  | -1 |
| 9   | 1  | -1 | -1 | -1 | -1 | 1  | -1 | -1 | 1  | 1  | 1  | -1 |
| 10  | 1  | -1 | -1 | -1 | 1  | -1 | 1  | -1 | -1 | -1 | -1 | -1 |
| 11  | 1  | -1 | 1  | 1  | -1 | -1 | -1 | 1  | -1 | 1  | 1  | 1  |
| 12  | 1  | -1 | 1  | 1  | 1  | 1  | 1  | 1  | 1  | -1 | -1 | 1  |
| 13  | 1  | 1  | -1 | -1 | -1 | 1  | 1  | 1  | -1 | 1  | -1 | 1  |
| 14  | 1  | 1  | -1 | 1  | 1  | -1 | -1 | 1  | 1  | 1  | -1 | -1 |
| 15  | 1  | 1  | 1  | -1 | 1  | 1  | -1 | -1 | -1 | -1 | 1  | 1  |
| 16  | 1  | 1  | 1  | 1  | -1 | -1 | 1  | -1 | 1  | -1 | 1  | -1 |

**TABLE 4.37 Recommended 16-Run 13-Factor No-Confounding Design**

| Run | A  | B  | C  | D  | E  | F  | G  | H  | J  | K  | L  | M  | N  |
|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1   | -1 | -1 | -1 | 1  | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  |
| 2   | -1 | -1 | 1  | -1 | -1 | -1 | -1 | -1 | 1  | 1  | -1 | -1 | 1  |
| 3   | -1 | -1 | 1  | -1 | 1  | 1  | 1  | 1  | 1  | -1 | 1  | -1 | -1 |
| 4   | -1 | -1 | 1  | 1  | -1 | 1  | 1  | 1  | -1 | 1  | -1 | 1  | -1 |
| 5   | -1 | 1  | -1 | -1 | -1 | 1  | -1 | -1 | -1 | -1 | 1  | -1 | -1 |
| 6   | -1 | 1  | -1 | -1 | 1  | 1  | 1  | -1 | -1 | 1  | -1 | 1  | 1  |
| 7   | -1 | 1  | -1 | 1  | -1 | -1 | 1  | 1  | 1  | -1 | 1  | 1  | 1  |
| 8   | -1 | 1  | 1  | 1  | 1  | -1 | -1 | -1 | 1  | -1 | -1 | 1  | -1 |
| 9   | 1  | -1 | -1 | -1 | -1 | -1 | 1  | -1 | 1  | 1  | 1  | 1  | -1 |
| 10  | 1  | -1 | -1 | -1 | 1  | -1 | -1 | 1  | -1 | -1 | -1 | 1  | -1 |
| 11  | 1  | -1 | -1 | 1  | 1  | 1  | 1  | -1 | 1  | -1 | -1 | -1 | 1  |
| 12  | 1  | -1 | 1  | 1  | -1 | 1  | -1 | -1 | -1 | -1 | 1  | 1  | 1  |
| 13  | 1  | 1  | -1 | 1  | -1 | 1  | -1 | 1  | 1  | 1  | -1 | -1 | -1 |
| 14  | 1  | 1  | 1  | -1 | -1 | -1 | 1  | 1  | -1 | -1 | -1 | -1 | 1  |
| 15  | 1  | 1  | 1  | -1 | 1  | 1  | -1 | 1  | 1  | 1  | 1  | 1  | 1  |
| 16  | 1  | 1  | 1  | 1  | 1  | -1 | 1  | -1 | -1 | 1  | 1  | -1 | -1 |

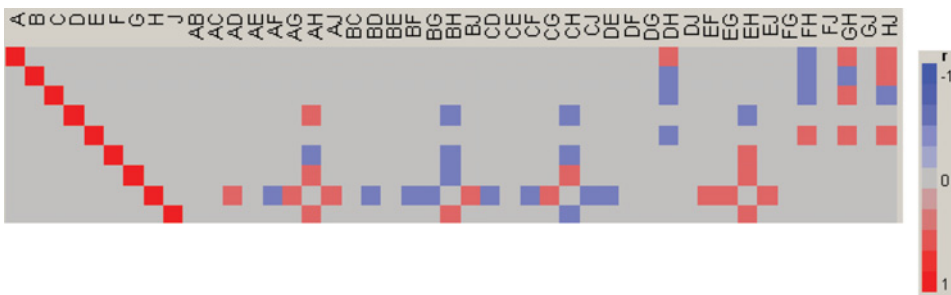
**TABLE 4.38 Recommended 16-Run 14-Factor No-Confounding Design**

| Run | A  | B  | C  | D  | E  | F  | G  | H  | J  | K  | L  | M  | N  | P  |
|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1   | -1 | -1 | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  | 1  | -1 | -1 | 1  |
| 2   | -1 | -1 | -1 | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  | 1  | -1 |
| 3   | -1 | -1 | 1  | -1 | -1 | 1  | -1 | 1  | 1  | 1  | -1 | -1 | 1  | 1  |
| 4   | -1 | -1 | 1  | 1  | 1  | 1  | 1  | -1 | -1 | -1 | 1  | -1 | 1  | -1 |
| 5   | -1 | 1  | -1 | -1 | -1 | 1  | 1  | -1 | 1  | -1 | 1  | 1  | -1 | 1  |
| 6   | -1 | 1  | -1 | 1  | 1  | 1  | -1 | 1  | -1 | 1  | -1 | 1  | -1 | -1 |
| 7   | -1 | 1  | 1  | -1 | -1 | -1 | -1 | 1  | -1 | -1 | 1  | 1  | 1  | -1 |
| 8   | -1 | 1  | 1  | 1  | 1  | -1 | -1 | -1 | 1  | -1 | -1 | -1 | -1 | 1  |
| 9   | 1  | -1 | -1 | -1 | 1  | 1  | -1 | -1 | -1 | -1 | -1 | 1  | 1  | 1  |
| 10  | 1  | -1 | -1 | 1  | -1 | 1  | -1 | 1  | 1  | -1 | 1  | -1 | -1 | -1 |
| 11  | 1  | -1 | 1  | -1 | 1  | -1 | -1 | -1 | 1  | 1  | 1  | 1  | -1 | -1 |
| 12  | 1  | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 | -1 | 1  | -1 | 1  |
| 13  | 1  | 1  | -1 | -1 | 1  | -1 | 1  | 1  | 1  | -1 | -1 | -1 | 1  | -1 |
| 14  | 1  | 1  | -1 | 1  | -1 | -1 | -1 | -1 | -1 | 1  | 1  | -1 | 1  | 1  |
| 15  | 1  | 1  | 1  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | -1 | -1 | -1 | -1 |
| 16  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |

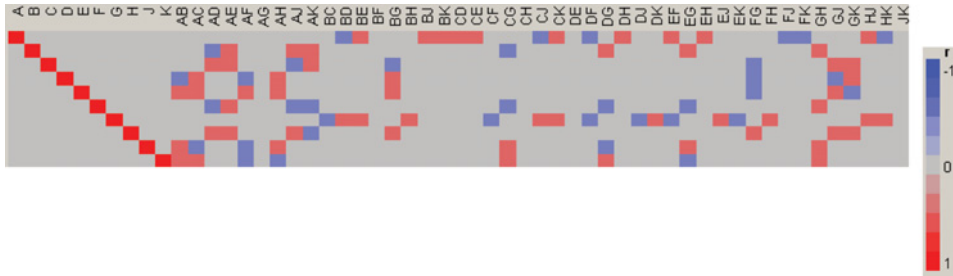
of the Hall designs and can also be found by the variance-bias trade-off optimization procedure developed by Jones and Nachtsheim (2011b). For more details including discussion of projection properties, see Jones, Shinde, and Montgomery (2015) and Montgomery (2013). The correlation matrices of the main effects with all two-factor interactions for the designs are shown in Figures 4.22 through 4.27. All recommended designs are first-order orthogonal (100% *D*-efficient) and the correlations between main effects and two-factor interactions are  $\pm 0.5$ .

**4.8.3 Analysis of Nonregular Fractional Factorial Designs**

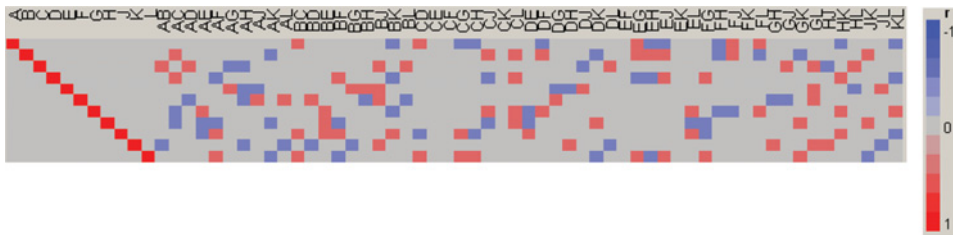
Forward selection regression can be a very useful approach for analyzing nonregular designs. There are also variations of the procedure that may be helpful in some situations.



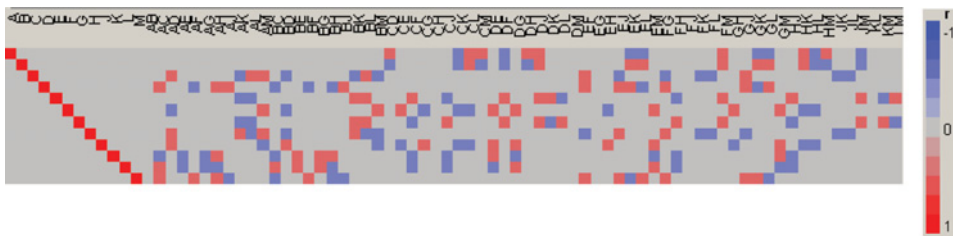
**Figure 4.22** Correlations of main effects and two-factor interactions for the no-confounding design for 9 factors in 16 runs.



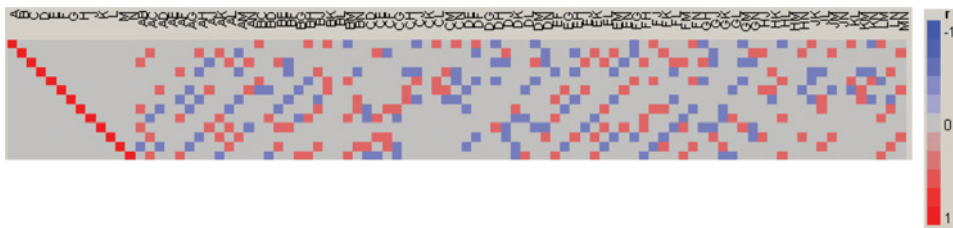
**Figure 4.23** Correlations of main effects and two-factor interactions for the no-confounding design for 10 factors in 16 runs.



**Figure 4.24** Correlations of main effects and two-factor interactions for the no-confounding design for 11 factors in 16 runs.

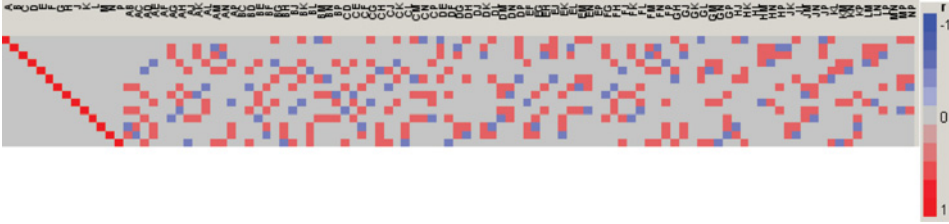


**Figure 4.25** Correlations of main effects and two-factor interactions for the no-confounding design for 12 factors in 16 runs.



**Figure 4.26** Correlations of main effects and two-factor interactions for the no-confounding design for 13 factors in 16 runs.





**Figure 4.27** Correlations of main effects and two-factor interactions for the no-confounding design for 14 factors in 16 runs.

Let's begin the discussion by identifying the types of models that may be of interest. We assume that main effects and two-factor interactions may be important and that higher-order interactions are negligible. Interactions may be **hierarchical**; that is, an interaction  $AB$  (say) may be in the model only if both of the main effects (here both  $A$  and  $B$ ) are also in the model. This situation occurs frequently in practice, so assuming hierarchy in analyzing data from a nonregular design is usually not a bad assumption. Interactions may also obey only a **heredity** principle (some authors say weak heredity); this means that  $AB$  can be in the model if either  $A$  or  $B$  is in the model. This situation also occurs fairly often, although not as often as hierarchy, but ignoring this possibility could result in the experimenter failing to identify all of the large interactions effects. Finally, there can be situations where an interaction such as  $AB$  is active but neither main effect  $A$  or  $B$  is active. This case is relatively uncommon.

There are several variations of forward selection regression that are useful. They are briefly described as follows:

1. Use forward selection without concern about model hierarchy. This can result in including too many interaction terms.
2. Use forward selection restricted to hierarchy. This means that if  $AB$  is selected for entry, then the entire group of terms  $A$ ,  $B$ , and  $AB$  are entered in the model if  $A$  and  $B$  are not already included.
3. Consider using larger-than-“usual”  $P$ -values for entering factors. Many stepwise regression computer programs have “default” values for entering factors, such as  $P = 0.05$ . This may be too restrictive. Values of 0.10 or even 0.15 may work better. The big danger in screening designs is not identifying important effects (type II errors), so type I errors are usually a lesser concern.
4. Use forward selection in two steps. First, select terms from all the main effects. Then run forward selection a second time using all two-factor interactions that satisfy the heredity assumption based on the main effects identified in step 1. You could also include any two-factor interactions that experience or process knowledge suggests should be considered.

Another alternative to forward stepwise regression is to consider some variation of all-possible-models regression. This is a procedure where we fit all possible regression models of particular sizes (such as all possible one-factor models, all-possible two-factor models, etc.) and use some criterion such as minimum mean square error, or restrict attention to hierarchical models, or models associated with a large increase in adjusted  $R^2$ , to narrow down the set of possible models for further consideration.

In some cases, nonregular designs have useful projection properties, and this could suggest an appropriate analysis. For example, the 16-run no-confounding designs for 6, 7, and 8 factors will support a model with all main effects and all two-factor interactions in any  $k = 4$  factors. So if up to four main effects appear large, we could analyze this design simply by fitting the main effects plus two-factor interaction model to the four apparently active effects. In such situations, it still may be useful to consider other possible interaction terms for inclusion in the final model.

#### 4.9 FRACTIONAL FACTORIAL SPLIT-PLOT DESIGNS

In Chapter 3 we introduced the **split-plot design** as a technique for dealing with hard-to-change versus easy-to-change factors in an experiment. Recall that the hard-to-change variables are changed less frequently in the experiment and placed in the whole plots, while the easy-to-change variables are still reset for each experimental run and are placed in the subplots. Factorial experiments with three or more factors in a split-plot structure often tend to be rather large experiments. On the other hand, the split-plots structure often makes it easier to conduct a larger experiment because we have reduced the number of times the hard-to-change factors need to be reset, and it can be quite easy and inexpensive to change the levels of the subplot variables.

As the number of factors in the experiment grows, the experimenter may consider using a fractional factorial experiment in the split-plot setting. As an illustration, consider an experiment involving five factors that potentially affect uniformity in a single-wafer plasma etching process. Three of the factors on the etching tool are relatively difficult to change from run to run:  $A$  = electrode gap,  $B$  = gas flow, and  $C$  = pressure. Two other factors are easy to change from run to run:  $D$  = time and  $E$  = RF (radio frequency) power. For subsequent clarity, we have used boldface letters to represent the easy-to-change factors. The experimenters want to use a fractional factorial split-plot (FFSP) design to investigate these five factors because the number of test wafers available is limited. The hard-to-change factors also indicate that a split-plot design should be considered to reduce the total number of changes required. The experimenters decide to use a  $2^{5-1}$  design with factors  $A$ ,  $B$ , and  $C$  in the whole plots and factors  $D$  and  $E$  in the subplots. The design generator is  $E = ABCD$ . This produces a 16-run fractional factorial with eight whole plots. Every whole plot contains one of the eight treatment combinations from a complete  $2^3$  factorial design in factors  $A$ ,  $B$ , and  $C$ . Each whole plot is divided into two subplots, with one of the treatment combinations for factors  $D$  and  $E$  in each subplot. The resulting FFSP design with measured responses of uniformity is shown in Table 4.39.

We will assume that all three-, four- and five-factor interactions are negligible. If this assumption is reasonable, then with 16 total observations the 15 effects (5 main effect and 10 two-way interactions) can be estimated. In the whole plots all hard-to-change factors,  $A$ ,  $B$ , and  $C$ , and their two-factor interactions,  $AB$ ,  $AC$ , and  $BC$ , can be estimated from the eight whole plots. There are seven degrees of freedom available for estimating whole plot effects, with the remaining degree of freedom from the whole plot terms being assigned to  $ABC = DE$ . There are eight degrees of freedom available for estimating subplot terms, which are for the two main effects,  $D$  and  $E$ , and all of the interactions involving a hard-to-change and easy-to-change factor ( $AD$ ,  $AE$ ,  $BD$ ,  $BE$ ,  $CD$ ,  $CE$ ). This illustrates one of the interesting trade-offs for using a fractional factorial split-plot design: Since we have a reduced number of runs for the experiment, and the defining relation involves both hard-to-change and

**TABLE 4.39 The Fractional Factorial Split-Plot Design for the Plasma Etching Experiment**

| Whole Plots | Whole-Plot Factors |          |          | Subplot Factors |          | Uniformity |
|-------------|--------------------|----------|----------|-----------------|----------|------------|
|             | <i>A</i>           | <i>B</i> | <i>C</i> | <i>D</i>        | <i>E</i> |            |
| 1           | –                  | –        | –        | –               | +        | 40.85      |
|             |                    |          |          | +               | –        | 41.07      |
| 2           | +                  | –        | –        | –               | –        | 35.67      |
|             |                    |          |          | +               | +        | 51.15      |
| 3           | –                  | +        | –        | –               | –        | 41.80      |
|             |                    |          |          | +               | +        | 37.01      |
| 4           | +                  | +        | –        | –               | +        | 91.09      |
|             |                    |          |          | +               | –        | 48.67      |
| 5           | –                  | –        | +        | –               | –        | 40.32      |
|             |                    |          |          | +               | +        | 43.34      |
| 6           | +                  | –        | +        | –               | +        | 62.46      |
|             |                    |          |          | +               | –        | 38.08      |
| 7           | –                  | +        | +        | –               | +        | 31.99      |
|             |                    |          |          | +               | –        | 41.03      |
| 8           | +                  | +        | +        | –               | –        | 70.31      |
|             |                    |          |          | +               | +        | 81.03      |

easy-to-change factors, care needs to be taken to determine which terms can be estimated with which degrees of freedom.

Each term in our model is confounded with one other effect. Therefore, both should be examined to determine how the effect will be estimated. If either of the two confounded effects involves only hard-to-change factors, then this term will need to be estimated with a whole-plot degree of freedom. For example, since  $ABC = DE$ , the two-factor interaction involving the two easy-to-change factors,  $DE$ , will be estimated as a whole-plot term. In general, any main effect or interaction that involves only hard-to-change factors will need to be compared to the whole-plot error. This also includes any terms that are aliased with higher-order terms that involve only whole-plot factors. Terms that involve any combination of factors that include at least one subplot factor in each term of the defining equation should be compared to the subplot error. See Montgomery (2013) for a more thorough discussion.

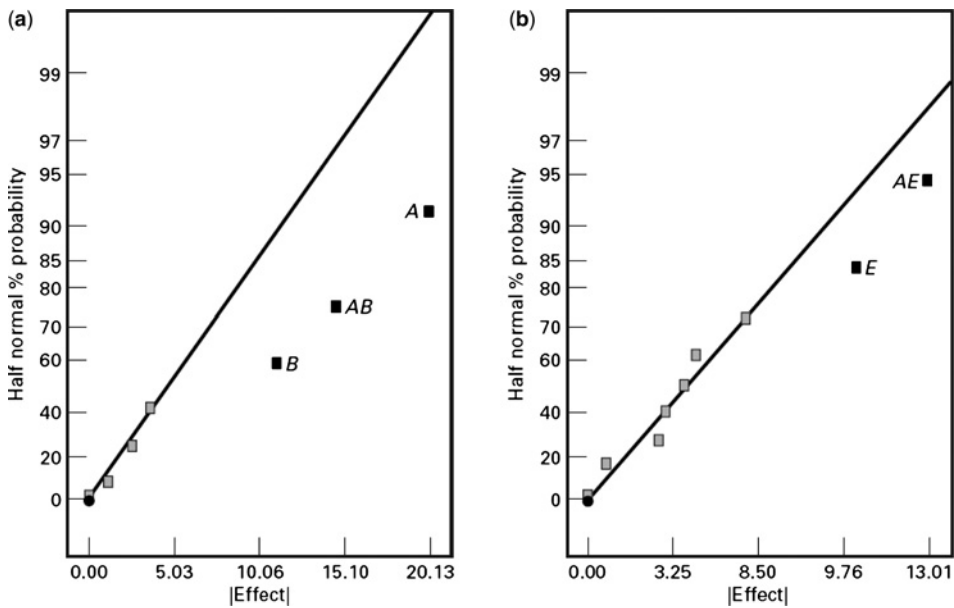
In this example, all of the two-way interactions involving a hard-to-change and easy-to-change factor are confounded with a three-way interaction involving both hard- and easy-to-change factors, and hence are estimated with subplot terms. For example,  $AD = BCE$  both involve at least one easy-to-change factor. Table 4.40 lists the effects and their estimates, separated into whole- and subplot terms. Since the terms of the model use all of the available degrees of freedom, the effects can be assessed via a normal probability plot. Because we have two error terms in the model, one for the whole-plot and one for the subplot, separate normal probability plots are required. Figure 4.28a is a half-normal probability plot of the effects estimates for the whole-plot terms,  $A$ ,  $B$ ,  $C$ ,  $AB$ ,  $AC$ ,  $BC$ , and  $DE$  ( $=ABC$ ). Notice that factors,  $A$ ,  $B$  and the  $AB$  interaction have large effects. Figure 4.28b is the half-normal probability plot of the subplot effects,  $D$ ,  $E$ ,  $AD$ ,  $AE$ ,  $BD$ ,  $BE$ ,  $CD$ , and  $CE$ . Only the main effect of  $E$  and  $AE$  interaction are large.

The design can be constructed using the JMP software package, by specifying the number of whole plots (8) and total observations (16) desired. The default design recommended by

**TABLE 4.40** Effects for Plasma Etching Experiment Separated into Whole Plot and Subplot Effects

| Term                  | Parameter Estimates | Type of Term |
|-----------------------|---------------------|--------------|
| Intercept             | 49.73875            |              |
| Gap ( <i>A</i> )      | 10.0625             | Whole        |
| Gas flow ( <i>B</i> ) | 5.6275              | Whole        |
| Pressure ( <i>C</i> ) | 1.325               | Whole        |
| Time ( <i>D</i> )     | -2.0725             | Subplot      |
| RF power ( <i>E</i> ) | 5.12625             | Subplot      |
| <i>AB</i>             | 7.34625             | Whole        |
| <i>AC</i>             | 1.83125             | Whole        |
| <i>AD</i>             | -3.00875            | Subplot      |
| <i>AE</i>             | 6.505               | Subplot      |
| <i>BC</i>             | -0.60125            | Whole        |
| <i>BD</i>             | -1.35875            | Subplot      |
| <i>BE</i>             | -0.2125             | Subplot      |
| <i>CD</i>             | 1.86625             | Subplot      |
| <i>CE</i>             | -1.485              | Subplot      |
| <i>DE</i>             | 0.34                | Whole        |

the software is the 32-run design shown in Table 4.41. This 32-run design is actually a full factorial split-plot design with eight whole plots and four subplots per whole plot, not a fractional factorial split-plot. It allows estimation of both the whole-plot and subplot error terms. It should be noted that the comparison between the two designs in Tables 4.39 and 4.41 involves considering the relative increase in cost from 16 to 32 runs, but since both



**Figure 4.28** Half-normal plots of the effects from the  $2^{5-1}$  fractional factorial split-plot design for the plasma etching experiment. (a) Whole-plot effects. (b) Subplot effects.

**TABLE 4.41** Default 32-Run Design from JMP for the Plasma Etching Experiment

| Whole Plots | Whole-Plot Factors |   |   | Subplot Factors |   |
|-------------|--------------------|---|---|-----------------|---|
|             | A                  | B | C | D               | E |
| 1           | -                  | - | - | -               | - |
|             |                    |   |   | -               | + |
|             |                    |   |   | +               | - |
| 2           | -                  | - | + | +               | + |
|             |                    |   |   | -               | - |
|             |                    |   |   | -               | + |
|             |                    |   |   | +               | - |
| 3           | -                  | + | - | +               | + |
|             |                    |   |   | -               | - |
|             |                    |   |   | -               | + |
|             |                    |   |   | +               | - |
| 4           | -                  | + | + | +               | + |
|             |                    |   |   | -               | - |
|             |                    |   |   | -               | + |
|             |                    |   |   | +               | - |
| 5           | +                  | - | - | +               | + |
|             |                    |   |   | -               | - |
|             |                    |   |   | -               | + |
|             |                    |   |   | +               | - |
| 6           | +                  | - | + | +               | + |
|             |                    |   |   | -               | - |
|             |                    |   |   | -               | + |
|             |                    |   |   | +               | - |
| 7           | +                  | + | - | +               | + |
|             |                    |   |   | -               | - |
|             |                    |   |   | -               | + |
|             |                    |   |   | +               | - |
| 8           | +                  | + | + | +               | + |
|             |                    |   |   | -               | - |
|             |                    |   |   | -               | + |
|             |                    |   |   | +               | - |
|             |                    |   |   | +               | + |

designs just involve eight whole plots, they both involve the same number of changes to the hard-to-change factors. Hence the actual cost of running the smaller design might not be much different than the total cost of the full factorial. In this case, the restriction on the number of test wafers available forces the choice of the FFSP.

#### 4.10 SUMMARY

This chapter has introduced the  $2^{k-p}$  fractional factorial design. We have emphasized the use of these designs in screening experiments to identify quickly and efficiently the subset of factors that are active and to provide some information on interaction. The projection property of these designs makes it possible in many cases to examine the active factors in

**TABLE 4.42 Useful Factorial and Fractional Factorial Designs from the  $2^{k-p}$  System**

| Design Type             | Number of Factors in Experiment |     |      |         |
|-------------------------|---------------------------------|-----|------|---------|
|                         | 4                               | 8   | 16   | 32 runs |
| Full factorial          | 2                               | 3   | 4    | 5       |
| Half-fraction           | 3                               | 4   | 5    | 6       |
| Resolution IV fraction  | —                               | 4   | 6–8  | 7–16    |
| Resolution III fraction | 3                               | 5–7 | 9–15 | 17–31   |

more detail. Sequential assembly of these designs via fold-over is a very effective way to gain additional information about interactions that an initial experiment may identify as possibly important.

In practice,  $2^{k-p}$  fractional factorial designs with  $N = 4, 8, 16,$  and  $32$  runs are highly useful. Table 4.42 summarizes these designs, identifying how many factors can be used with each design to obtain various types of screening experiments. For example, the 16-run design is a full factorial for 4 factors, a one-half fraction for 5 factors, resolution IV fractional design for 6–8 factors, and a resolution III fraction for 9–15 factors. All of these designs may be constructed using the methods discussed in this chapter. Montgomery (2013) gives the alias relationships for many of these designs.

We also introduced nonregular designs and analysis strategies for screening for important factors. These represent an additional set of tools for experiments with large numbers of factors.

## EXERCISES

- 4.1 Suppose that in the chemical process development experiment described in Exercise 3.6, it was only possible to run a one-half fraction of the  $2^3$  design. Construct the design and perform the statistical analysis by selecting the relevant runs.
- 4.2 Suppose that in Exercise 3.9, only a one-half fraction of the  $2^4$  design could be run. Construct the design and perform the analysis by selecting the relevant runs.
- 4.3 Consider the plasma etch experiment described in Exercise 3.10. Suppose that only a one-half fraction of the design could be run. Set up the design and analyze the data.
- 4.4 Example 4.2 describes a process improvement study in the manufacture of an integrated circuit. Suppose that only eight runs could be made in this process. Set up an appropriate  $2^{5-2}$  design and find the alias structure. Use the data from Example 4.2 as the observations in this design, and estimate the factor effects. What conclusions can you draw?
- 4.5 **Continuation of Exercise 4.4.** Suppose you have made the eight runs in the  $2^{5-2}$  design in Exercise 4.4. What additional runs would be required to identify the factor effects that are of interest? What are the alias relationships in the combined design?
- 4.6 R. D. Snee (“Experimenting with a Large Number of Variables,” in *Experiments in Industry: Design, Analysis and Interpretation of Results*, by R. D. Snee, L. B. Hare, and J. B. Trout, editors, ASQC, 1985) describes an experiment in which a  $2^{5-1}$

design with  $I = ABCDE$  was used to investigate that effects of five factors on the color of a chemical product. The factors are  $A =$  solvent/reactant,  $B =$  catalyst/reactant,  $C =$  temperature,  $D =$  reactant purity, and  $D =$  reactant pH. The results obtained were as follows:

$$\begin{array}{ll} e = -0.63 & d = 6.79 \\ a = 2.51 & ade = 5.47 \\ b = -2.68 & bde = 3.45 \\ abe = 1.66 & abd = 5.68 \\ c = 2.06 & cde = 5.22 \\ ace = 1.22 & acd = 4.38 \\ bce = -2.09 & bcd = 4.30 \\ abc = 1.93 & abcde = 4.05 \end{array}$$

- (a) Prepare a normal probability plot of the effects. Which effects seem active?
- (b) Calculate the residuals. Construct a normal probability plot of the residuals and plot the residuals versus the fitted values. Comment on the plots.
- (c) If any factors are negligible, collapse the  $2^{5-1}$  design into a full factorial in the active factors. Comment on the resulting design, and interpret the results.
- 4.7 An article in the *Journal of Quality Technology* (Vol. 17, 1985, pp. 198–206) describes the use of a replicated fractional factorial to investigate the effect of five factors on the free height of leaf springs used in a automotive application. The factors are  $A =$  furnace temperature,  $B =$  heating time,  $C =$  transfer time,  $D =$  hold-down time, and  $E =$  quench oil temperature. The data are shown in Table E4.1.
- (a) Write out the alias structure for this design. What is the resolution of this design?
- (b) Analyze the data. What factors influence the mean free height?
- (c) Calculate the range and standard deviation of the free height for each run. Is there any indication that any of these factors affects variability in the free height?
- (d) Analyze the residuals from this experiment, and comment on your findings.

**TABLE E4.1 The Leot Spring Experiment from Exercise 4.7**

| $A$ | $B$ | $C$ | $D$ | $E$ | Free Height |      |      |
|-----|-----|-----|-----|-----|-------------|------|------|
| –   | –   | –   | –   | –   | 7.78        | 7.78 | 7.81 |
| +   | –   | –   | +   | –   | 8.15        | 8.18 | 7.88 |
| –   | +   | –   | +   | –   | 7.50        | 7.56 | 7.50 |
| +   | +   | –   | –   | –   | 7.59        | 7.56 | 7.75 |
| –   | –   | +   | +   | –   | 7.54        | 8.00 | 7.88 |
| +   | –   | +   | –   | –   | 7.69        | 8.09 | 8.06 |
| –   | +   | +   | –   | –   | 7.56        | 7.52 | 7.44 |
| +   | +   | +   | +   | –   | 7.56        | 7.81 | 7.69 |
| –   | –   | –   | –   | +   | 7.50        | 7.25 | 7.12 |
| +   | –   | –   | +   | +   | 7.88        | 7.88 | 7.44 |
| –   | +   | –   | +   | +   | 7.50        | 7.56 | 7.50 |
| +   | +   | –   | –   | +   | 7.63        | 7.75 | 7.56 |
| –   | –   | +   | +   | +   | 7.32        | 7.44 | 7.44 |
| +   | –   | +   | –   | +   | 7.56        | 7.69 | 7.62 |
| –   | +   | +   | –   | +   | 7.18        | 7.18 | 7.25 |
| +   | +   | +   | +   | +   | 7.81        | 7.50 | 7.59 |

- (e) Is this the best possible design for five factors in 16 runs? Specifically, can you find a fractional design for five factors in 16 runs with a higher resolution than this one?
- 4.8** Reconsider the experiment in Exercise 4.6. Suppose that along with the original 16 runs the experimenter had run four center points with the following observed responses: 5.20, 4.98, 5.26, and 5.40.
- (a) Use the center points to obtain an estimate of pure error.  
 (b) Test for curvature in the response function and lack of fit.  
 (c) What type of model seems appropriate for this purpose?  
 (d) Can you fit the model in part (c) using the data from this experiment?
- 4.9** An article in *Industrial and Engineering Chemistry* (“More on Planning Experiments to Increase Research Efficiency,” 1970, pp. 60–65) uses a  $2^{5-2}$  design to investigate the effect of  $A$  = condensation temperature,  $B$  = amount of material 1,  $C$  = solvent volume,  $D$  = condensation time, and  $E$  = amount of material 2 on yield. The results obtained are as follows:

$$\begin{array}{cccc} e = 23.2 & ad = 16.9 & cd = 23.8 & bde = 16.8 \\ ab = 15.5 & bc = 16.2 & ace = 23.4 & abcde = 18.1 \end{array}$$

- (a) Verify that the design generators used were  $I = ACE$  and  $I = BDE$ .  
 (b) Write down the complete defining relation and the aliases for this design.  
 (c) Estimate the main effects.  
 (d) Prepare an analysis of variance table. Verify that the  $AB$  and  $AD$  interactions are available to use as error.  
 (e) Plot the residuals versus the fitted values. Also construct a normal probability plot of the residuals. Comment on the results.
- 4.10** Reconsider the experiment in Exercise 4.9. Suppose that four center points were run, and that the responses are as follows: 20.1, 20.9, 19.8, and 20.4.
- (a) Use the center points to obtain an estimate of pure error.  
 (b) Test for lack of fit and curvature. What conclusions can you draw about the type of model required for yield in this process?
- 4.11** An industrial engineer is conducting an experiment using a Monte Carlo simulation model of an inventory system. The independent variables in her model are the order quantity ( $A$ ), the reorder point ( $B$ ), the setup cost ( $C$ ), the backorder cost ( $D$ ), and the carrying cost rate ( $E$ ). The response variable is average annual cost. To conserve computer time, she decides to investigate these factors using a  $2^{5-2}_{III}$  design with  $I = ABD$  and  $I = BCE$ . The results she obtains are  $de = 95$ ,  $ae = 134$ ,  $b = 158$ ,  $abd = 190$ ,  $cd = 92$ ,  $ac = 187$ ,  $bce = 155$ , and  $abcde = 185$ .
- (a) Verify that the treatment combinations given are correct. Estimate the effects, assuming three-factor and higher interactions are negligible.  
 (b) Suppose that a second fraction is added to the first. The runs in this new design are  $ade = 136$ ,  $e = 93$ ,  $ab = 187$ ,  $bd = 153$ ,  $acd = 139$ ,  $c = 99$ ,  $abce = 191$ , and  $bcd = 150$ . How was this second fraction obtained? Add these runs to the original fraction, and estimate the effects.



(c) Suppose that the fraction  $abc = 189$ ,  $ce = 96$ ,  $bcd = 154$ ,  $acde = 135$ ,  $abe = 193$ ,  $bde = 152$ ,  $ad = 137$ , and  $(1) = 98$  was run. How was this fraction obtained? Add these data to the original fraction, and estimate the effects.

**4.12** Carbon anodes used in a smelting process are baked in a ring furnace. An experiment is run in the furnace to determine which factors influence the weight of packing material that is stuck to the anodes after baking. Six variables are of interest, each at two levels:  $A$  = pitch/fines ratio (0.45, 0.55);  $B$  = packing material type (1, 2);  $C$  = packing material temperature (ambient, 325°C);  $D$  = flue location (inside, outside);  $E$  = pit temperature (ambient, 195°C); and  $F$  = delay time before packing (zero, 24 hours). A  $2^{6-3}$  design is run, and three replicates are obtained at each of the design points. The weight of packing material stuck to the anodes is measured in grams. The data in run order are as follows:  $abd = (984, 826, 936)$ ;  $abcdef = (1275, 976, 1457)$ ;  $be = (1217, 1201, 890)$ ;  $af = (1474, 1164, 1541)$ ;  $def = (1320, 1156, 913)$ ;  $cd = (765, 705, 821)$ ;  $ace = (1338, 1254, 1294)$ ; and  $bcf = (1325, 1299, 1253)$ . We wish to minimize the amount of stuck packing material.

- Verify that the eight runs correspond to a  $2^{6-3}_{III}$  design. What is the alias structure?
- Use the average weight as a response. What factors appear to be influential?
- Use the range of the weights as a response. What factors appear to be influential?
- What recommendations would you make to the process engineers?

**4.13** A 16-run experiment was performed in a semiconductor manufacturing plant to study the effects of six factors on the curvature, or **camber**, of the substrate devices produced. The six variables and their levels are shown in Table E4.2. Each run was replicated four times, and a camber measurement was taken on the substrate. The data are shown in Table E4.3.

**TABLE E4.2 Factors Levels for the Experiment in Exercise 4.13**

| Run | Lamination Temperature (°C) | Lamination Time (sec) | Lamination Pressure (ton) | Firing Temperature (°C) | Firing Cycle Time (hr) | Firing Dew Point (°C) |
|-----|-----------------------------|-----------------------|---------------------------|-------------------------|------------------------|-----------------------|
| 1   | 55                          | 10                    | 5                         | 1580                    | 17.5                   | 20                    |
| 2   | 75                          | 10                    | 5                         | 1580                    | 29                     | 26                    |
| 3   | 55                          | 25                    | 5                         | 1580                    | 29                     | 20                    |
| 4   | 75                          | 25                    | 5                         | 1580                    | 17.5                   | 26                    |
| 5   | 55                          | 10                    | 10                        | 1580                    | 29                     | 26                    |
| 6   | 75                          | 10                    | 10                        | 1580                    | 17.5                   | 20                    |
| 7   | 55                          | 25                    | 10                        | 1580                    | 17.5                   | 26                    |
| 8   | 75                          | 25                    | 10                        | 1580                    | 29                     | 20                    |
| 9   | 55                          | 10                    | 5                         | 1620                    | 17.5                   | 26                    |
| 10  | 75                          | 10                    | 5                         | 1620                    | 29                     | 20                    |
| 11  | 55                          | 25                    | 5                         | 1620                    | 29                     | 26                    |
| 12  | 75                          | 25                    | 5                         | 1620                    | 17.5                   | 20                    |
| 13  | 55                          | 10                    | 10                        | 1620                    | 29                     | 20                    |
| 14  | 75                          | 10                    | 10                        | 1620                    | 17.5                   | 26                    |
| 15  | 55                          | 25                    | 10                        | 1620                    | 17.5                   | 20                    |
| 16  | 75                          | 25                    | 10                        | 1620                    | 29                     | 26                    |

**TABLE E4.3 Data for the Experiment in Exercise 4.13**

| Run | Camber for Replicate (in./in.) |        |        |        | Total<br>( $10^{-4}$ in./in.) | Mean<br>( $10^{-4}$ in./in.) | Standard<br>Deviation |
|-----|--------------------------------|--------|--------|--------|-------------------------------|------------------------------|-----------------------|
|     | 1                              | 2      | 3      | 4      |                               |                              |                       |
| 1   | 0.0167                         | 0.0128 | 0.0149 | 0.0185 | 629                           | 157.25                       | 24.418                |
| 2   | 0.0062                         | 0.0066 | 0.0044 | 0.0020 | 192                           | 48.00                        | 20.976                |
| 3   | 0.0041                         | 0.0043 | 0.0042 | 0.0050 | 176                           | 44.00                        | 4.0825                |
| 4   | 0.0073                         | 0.0071 | 0.0039 | 0.0030 | 223                           | 55.75                        | 25.025                |
| 5   | 0.0047                         | 0.0047 | 0.0040 | 0.0089 | 223                           | 55.75                        | 22.410                |
| 6   | 0.0219                         | 0.0258 | 0.0147 | 0.0296 | 920                           | 230.00                       | 63.639                |
| 7   | 0.0121                         | 0.0090 | 0.0092 | 0.0086 | 389                           | 97.25                        | 16.029                |
| 8   | 0.0255                         | 0.0250 | 0.0226 | 0.0169 | 900                           | 225.00                       | 39.42                 |
| 9   | 0.0032                         | 0.0023 | 0.0077 | 0.0069 | 201                           | 50.25                        | 26.725                |
| 10  | 0.0078                         | 0.0158 | 0.0060 | 0.0045 | 341                           | 85.25                        | 50.341                |
| 11  | 0.0043                         | 0.0027 | 0.0028 | 0.0028 | 126                           | 31.50                        | 7.681                 |
| 12  | 0.0186                         | 0.0137 | 0.0158 | 0.0159 | 640                           | 160.00                       | 20.083                |
| 13  | 0.0110                         | 0.0086 | 0.0101 | 0.0158 | 455                           | 113.75                       | 31.12                 |
| 14  | 0.0065                         | 0.0109 | 0.0126 | 0.0071 | 371                           | 92.75                        | 29.51                 |
| 15  | 0.0155                         | 0.0158 | 0.0145 | 0.0145 | 603                           | 150.75                       | 6.75                  |
| 16  | 0.0093                         | 0.0124 | 0.0110 | 0.0133 | 460                           | 115.00                       | 17.45                 |

- (a) What type of design did the experimenters use?
- (b) What are the alias relationships in this design?
- (c) Do any of the process variables affect average camber?
- (d) Do any of the process variables affect the variability in camber measurements?
- (e) If it is important to reduce mean camber as much as possible, what recommendations would you make?
- (f) Fit an appropriate model for both mean camber and the standard deviation of camber. If we would like to reduce the variability in camber while keeping the mean camber as close as possible to  $100 \times 10^{-4}$  in./in., what recommendations would you make?
- 4.14** A spin coater is used to apply photoresist to a silicon wafer. This operation usually occurs early in the semiconductor manufacturing process, and the average coating thickness and the variability in the coating thickness have important effects on downstream manufacturing steps. Six variables are used in the experiment. The variables and their high and low levels are shown in Table E4.4. The experimenter decides to use a  $2^{6-1}$  design, and to make three readings on resist thickness on each test wafer. The data are shown in Table E4.5.
- (a) Verify that this is a  $2^{6-1}$  design. Discuss the alias relationships in this design.
- (b) What factors appear to affect average resist thickness?
- (c) Since the volume of resist applied has little effect on average thickness, does this have any important practical implications for the process engineers?
- (d) Project this design into a smaller design involving only significant factors. Display the result graphically. Does this aid in interpretation?

**TABLE E4.4 Factors and Levels for the Experiment in Exercise 4.14**

| Factor                   | Low Level | High Level |
|--------------------------|-----------|------------|
| Final spin speed         | 7300 rpm  | 6650 rpm   |
| Acceleration rate        | 5         | 20         |
| Volume of resist applied | 3 cc      | 5 cc       |
| Time of spin             | 14 s      | 6 s        |
| Resist batch variation   | Batch 1   | Batch 2    |
| Exhaust pressure         | Cover off | Cover on   |

**TABLE E4.5 Data for the Experiment in Exercise 4.14**

| Run | A Volume | B Batch | C Time | D Speed | E Acc. | F Cover | Resist Thickness |        |       |         |
|-----|----------|---------|--------|---------|--------|---------|------------------|--------|-------|---------|
|     |          |         |        |         |        |         | Left             | Center | Right | Average |
| 1   | 5        | Batch 2 | 14     | 7350    | 5      | Off     | 4531             | 4531   | 4515  | 4525.7  |
| 2   | 5        | Batch 1 | 6      | 7350    | 5      | Off     | 4446             | 4464   | 4428  | 4446    |
| 3   | 3        | Batch 1 | 6      | 6650    | 5      | Off     | 4452             | 4490   | 4452  | 4464.7  |
| 4   | 3        | Batch 2 | 14     | 7350    | 20     | Off     | 4316             | 4328   | 4308  | 4317.3  |
| 5   | 3        | Batch 1 | 14     | 7350    | 5      | Off     | 4307             | 4295   | 4289  | 4297    |
| 6   | 5        | Batch 1 | 6      | 6650    | 20     | Off     | 4470             | 4492   | 4495  | 4485.7  |
| 7   | 3        | Batch 1 | 6      | 7350    | 5      | On      | 4496             | 4502   | 4482  | 4493.3  |
| 8   | 5        | Batch 2 | 14     | 6650    | 20     | Off     | 4542             | 4547   | 4538  | 4542.3  |
| 9   | 5        | Batch 1 | 14     | 6650    | 5      | Off     | 4621             | 4643   | 4613  | 4625.7  |
| 10  | 3        | Batch 1 | 14     | 6650    | 5      | On      | 4653             | 4670   | 4645  | 4656    |
| 11  | 3        | Batch 2 | 14     | 6650    | 20     | On      | 4480             | 4486   | 4470  | 4478.7  |
| 12  | 3        | Batch 1 | 6      | 7350    | 20     | Off     | 4221             | 4233   | 4217  | 4223.7  |
| 13  | 5        | Batch 1 | 6      | 6650    | 5      | On      | 4620             | 4641   | 4619  | 4626.7  |
| 14  | 3        | Batch 1 | 6      | 6650    | 20     | On      | 4455             | 4480   | 4466  | 4467    |
| 15  | 5        | Batch 2 | 14     | 7350    | 20     | On      | 4255             | 4288   | 4243  | 4262    |
| 16  | 5        | Batch 2 | 6      | 7350    | 5      | On      | 4490             | 4534   | 4523  | 4515.7  |
| 17  | 3        | Batch 2 | 14     | 7350    | 5      | On      | 4514             | 4551   | 4540  | 4535    |
| 18  | 3        | Batch 1 | 14     | 6650    | 20     | Off     | 4494             | 4503   | 4496  | 4497.7  |
| 19  | 5        | Batch 2 | 6      | 7350    | 20     | Off     | 4293             | 4306   | 4302  | 4300.3  |
| 20  | 3        | Batch 2 | 6      | 7350    | 5      | Off     | 4534             | 4545   | 4512  | 4530.3  |
| 21  | 5        | Batch 1 | 14     | 6650    | 20     | On      | 4460             | 4457   | 4436  | 4451    |
| 22  | 3        | Batch 2 | 6      | 6650    | 5      | On      | 4650             | 4688   | 4656  | 4664.7  |
| 23  | 5        | Batch 1 | 14     | 7350    | 20     | Off     | 4231             | 4244   | 4230  | 4235    |
| 24  | 3        | Batch 2 | 6      | 7350    | 20     | On      | 4225             | 4228   | 4208  | 4220.3  |
| 25  | 5        | Batch 1 | 14     | 7350    | 5      | On      | 4381             | 4391   | 4376  | 4382.7  |
| 26  | 3        | Batch 2 | 6      | 6650    | 20     | Off     | 4533             | 4521   | 4511  | 4521.7  |
| 27  | 3        | Batch 1 | 14     | 7350    | 20     | On      | 4194             | 4230   | 4172  | 4198.7  |
| 28  | 5        | Batch 2 | 6      | 6650    | 5      | Off     | 4666             | 4695   | 4672  | 4677.7  |
| 29  | 5        | Batch 1 | 6      | 7350    | 20     | On      | 4180             | 4213   | 4197  | 4196.7  |
| 30  | 5        | Batch 2 | 6      | 6650    | 20     | On      | 4465             | 4496   | 4463  | 4474.7  |
| 31  | 5        | Batch 2 | 14     | 6650    | 5      | On      | 4653             | 4685   | 4665  | 4667.7  |
| 32  | 3        | Batch 2 | 14     | 6650    | 5      | Off     | 4683             | 4712   | 4677  | 4690.7  |

- (e) Use the range of resist thickness as a response variable. Is there any indication that any of these factors affect the variability in resist thickness?
- (f) Where would you recommend that the process engineers run the process?
- 4.15** Consider the leaf spring experiment in Exercise 4.7. Suppose that factor  $E$  (quench oil temperature) is very difficult to control during manufacturing. Where would you set factors  $A$ ,  $B$ ,  $C$ , and  $D$  to reduce variability in the free height as much as possible regardless of the quench oil temperature used?
- 4.16** Construct a  $2^{7-2}$  design by choosing two four-factor interactions as the independent generators. Write down the complete alias structure for this design. Outline the analysis of variance table. What is the resolution of this design?
- 4.17** Consider the  $2^5$  design in Exercise 3.12. Suppose that only a one-half fraction could be run. Furthermore, two days were required to take the 16 observations, and it was necessary to confound the  $2^{5-1}$  design in two blocks. Construct the design and analyze the data.
- 4.18** Analyze the data in the first replicate of Exercise 3.9 as if it came from a  $2^{4-1}_{IV}$  design with  $I = ABCD$ . Project the design into a full factorial in the subset of the original four factors that appear to be significant.
- 4.19** Repeat Exercise 4.18 using  $I = -ABCD$ . Does use of the alternate fraction change your interpretation of the data?
- 4.20** Project the  $2^{4-1}_{IV}$  design in Example 4.1 into two replicates of a  $2^2$  design in the factors  $A$  and  $B$ . Analyze the data and draw conclusions.
- 4.21** Construct a  $2^{6-3}_{III}$  design. Determine the effects that may be estimated if a second fraction of this design is run with all signs reversed.
- 4.22** Consider the  $2^{6-3}_{III}$  design in Exercise 4.21. Determine the effects that may be estimated if a second fraction of this design is run with the signs for factor  $A$  reversed.
- 4.23** Fold over the  $2^{7-3}_{III}$  design in Table 4.13. Verify that the resulting design is a  $2^{8-4}_{IV}$  design. Is this a minimal design?
- 4.24** Fold over a  $2^{5-2}_{III}$  design. Verify that the resulting design is a  $2^{6-2}_{IV}$  design. Compare this design with the  $2^{6-2}_{IV}$  design in Table 4.14.
- 4.25** Suppose that you have run a  $2^{4-1}_{IV}$  design with  $I = ABCD$ , and that you have fit the model  $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \varepsilon$ .
- (a) Assume that the true model for the system is  $E(y) = \beta_0 + \sum_{i=1}^4 \beta_i x_i + \beta_{12}x_1x_2$ . Are the least squares estimates of  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$ , and  $\beta_4$  in the fitted model unbiased? Is the result you obtained an “intuitive” one?
- (b) Assume that the true model is  $E(y) = \beta_0 + \sum_{i=1}^4 \beta_i x_i + \beta_{12}x_1x_2 + \beta_{11}x_1^2 + \beta_{22}x_2^2$ . Are the least squares estimates of  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$ , and  $\beta_4$  biased estimates?
- 4.26** Reconsider the experiment described in Exercise 4.13. Suppose that “firing temperature” is hard to change. Construct an appropriate fractional factorial split-plot design for this problem.

- 4.27** Reconsider the spin coating experiment from Exercise 4.14 with the factor “resist batch variation” as hard to change. Construct an appropriate fractional factorial split-plot design for this problem.
- 4.28** Consider the  $2^5$  factorial experiment given in Exercise 3.38. Suppose that only 16 runs could be made. Set up an appropriate  $2^{5-1}$  design for this problem and use only the data from those 16 runs in your analysis. Compare your results with those obtained from the full factorial in Exercise 3.38.
- 4.29** Reconsider the situation described in Exercise 4.28. Suppose that only 8 runs could be made in one time period. Set up the fractional factorial in two blocks of 8 runs each. Analyze the data from these 16 runs and comment on the results.
- 4.30** Consider the  $2^5$  factorial experiment given in Exercise 3.38. Suppose that only 8 runs could be made. Set up an appropriate  $2^{5-2}$  design for this problem and use only the data from those 8 runs in your analysis. Compare your results with those obtained from the full factorial in Exercise 3.38.
- 4.31** An article by L. B. Hare (“In the Soup: A Case Study to Identify Contributors to Filling Variability,” *Journal of Quality Technology*, Vol. 20, pp. 36–43) describes a factorial experiment used to study the filling variability of dry soup mix packages. The factors are  $A$  = number of mixing ports through which the vegetable oil was added (1, 2),  $B$  = temperature surrounding the mixer (cooled, ambient),  $C$  = mixing time (60, 80 sec),  $D$  = batch weight (1500, 2000 lb), and  $E$  = number of days of delay between mixing and packaging (1, 7). Between 125 and 150 packages of soup were sampled over an 8-hr period for each run in the design, and the standard deviation of package weight was used as the response variable. The design and resulting data are shown in Table E4.6

**TABLE E4.6 The Soup Experiment**

| Std. Order | $A$<br>Mixer Ports | $B$<br>Temp. | $C$<br>Time | $D$<br>Batch Weight | $E$<br>Delay | $y$<br>Std. Dev |
|------------|--------------------|--------------|-------------|---------------------|--------------|-----------------|
| 1          | –                  | –            | –           | –                   | –            | 1.13            |
| 2          | +                  | –            | –           | –                   | +            | 1.25            |
| 3          | –                  | +            | –           | –                   | +            | 0.97            |
| 4          | +                  | +            | –           | –                   | –            | 1.70            |
| 5          | –                  | –            | +           | –                   | +            | 1.47            |
| 6          | +                  | –            | +           | –                   | –            | 1.28            |
| 7          | –                  | +            | +           | –                   | –            | 1.18            |
| 8          | +                  | +            | +           | –                   | +            | 0.98            |
| 9          | –                  | –            | –           | +                   | +            | 0.78            |
| 10         | +                  | –            | –           | +                   | –            | 1.36            |
| 11         | –                  | +            | –           | +                   | –            | 1.85            |
| 12         | +                  | +            | –           | +                   | +            | 0.62            |
| 13         | –                  | –            | +           | +                   | –            | 1.09            |
| 14         | +                  | –            | +           | +                   | +            | 1.10            |
| 15         | –                  | +            | +           | +                   | +            | 0.76            |
| 16         | +                  | +            | +           | +                   | –            | 2.10            |

- (a) What is the generator for this design?
  - (b) What is the resolution of this design?
  - (c) Estimate the factor effects. Which effects are large? Form an appropriate model for the response.
  - (d) Does a residual analysis indicate any problems with the underlying assumptions?
  - (e) Draw conclusions about this filling process.
- 4.32** A 16-run fractional factorial experiment in 10 factors on sand-casting of engine manifolds was conducted by engineers at the Essex Aluminum Plant of the Ford Motor Company and described in the article “Evaporative Cast Process 3.0 Liter Intake Manifold Poor Sandfill Study,” by D. Becknell (*Fourth Symposium on Taguchi Methods*, American Supplier Institute, Dearborn, MI, 1986, pp. 120–130). The purpose was to determine which of 10 factors has an effect on the proportion of defective castings. The design and the resulting proportion of nondefective castings observed on each run are shown in Table E4.7. This is a resolution III fraction with generators  $E = CD, F = BD, G = BC, H = AC, J = AB$ , and  $K = ABC$ . Assume that the number of castings made at each run in the design is 1000.
- (a) Find the defining relation and the alias relationships in this design.
  - (b) Estimate the factor effects and use a normal probability plot to tentatively identify the important factors.
  - (c) Fit an appropriate model using the factors identified in part (b) above.
  - (d) Plot the residuals from this model versus the predicted proportion of nondefective castings. Also prepare a normal probability plot of the residuals. Comment on the adequacy of these plots.
  - (e) In part (d) you should have noticed an indication that the variance of the response is not constant. (Considering that the response is a proportion, you should have

**TABLE E4.7 The Sand-Casting Experiment in Exercise 4.32**

| Run | A | B | C | D | E | F | G | H | J | K | $\hat{p}$ | Aresin<br>$\sqrt{\hat{p}}$ | F&T's<br>Modification |
|-----|---|---|---|---|---|---|---|---|---|---|-----------|----------------------------|-----------------------|
| 1   | - | - | - | - | + | + | + | + | + | - | 0.958     | 1.364                      | 1.363                 |
| 2   | + | - | - | - | + | + | + | - | - | + | 1.000     | 1.571                      | 1.555                 |
| 3   | - | + | - | - | + | - | - | + | - | + | 0.977     | 1.419                      | 1.417                 |
| 4   | + | + | - | - | + | - | - | - | + | - | 0.775     | 1.077                      | 1.076                 |
| 5   | - | - | + | - | - | + | - | - | + | + | 0.958     | 1.364                      | 1.363                 |
| 6   | + | - | + | - | - | + | - | + | - | - | 0.958     | 1.364                      | 1.363                 |
| 7   | - | + | + | - | - | - | + | - | - | - | 0.813     | 1.124                      | 1.123                 |
| 8   | + | + | + | - | - | - | + | + | + | + | 0.906     | 1.259                      | 1.259                 |
| 9   | - | - | - | + | - | - | + | + | + | - | 0.679     | 0.969                      | 0.968                 |
| 10  | + | - | - | + | - | - | + | - | - | + | 0.781     | 1.081                      | 1.083                 |
| 11  | - | + | - | + | - | + | - | + | - | + | 1.000     | 1.571                      | 1.556                 |
| 12  | + | + | - | + | - | + | - | - | + | - | 0.896     | 1.241                      | 1.242                 |
| 13  | - | - | + | + | + | - | - | - | + | + | 0.958     | 1.364                      | 1.363                 |
| 14  | + | - | + | + | + | - | - | + | - | - | 0.818     | 1.130                      | 1.130                 |
| 15  | - | + | + | + | + | + | + | - | - | - | 0.841     | 1.161                      | 1.160                 |
| 16  | + | + | + | + | + | + | + | + | + | + | 0.955     | 1.357                      | 1,356                 |

expected this.) The previous table also shows a transformation on  $\hat{p}$ , the arcsin square root, which is a widely used variance stabilizing transformation for proportion data. Repeat parts (a) through (d) above using the transformed response and comment on your results. Specifically, are the residual plots improved?

(f) There is a modification to the arcsin square root transformation, proposed by Freeman and Tukey

(“Transformations Related to the Angular and the Square Root,” *Annals of Mathematical Statistics*, Vol. 21, 1950, pp. 607–611), that improves its performance in the tails. F&T’s modification is

$$[\arcsin \sqrt{n\hat{p}/(n+1)} + \arcsin \sqrt{(n\hat{p}+1)/(n+1)}]/2$$

Rework parts (a) through (d) using this transformation and comment on the results. (For an interesting discussion and analysis of this experiment, refer to “Analysis of Factorial Experiments with Defects or Defectives as the Response,” by S. Bisgaard and H. T. Fuller, *Quality Engineering*, Vol. 7, 1994–1995, pp. 429–443.)

**4.33** An experimenter has conducted a  $2^{5-1}$  fractional factorial experiment. The test matrix is shown in Table E4.8.

- (a) What is the resolution of the design?  
 (b) What are the design generators that have been used to construct this design? What is the complete defining relation?  
 (c) Write down the aliases, showing only main effects and two-factor interactions.

**TABLE E4.8** The  $2^{5-1}$  Design for Exercise 4.33

| Std | A  | B  | C  | D  | E  | Block |
|-----|----|----|----|----|----|-------|
| 1   | -1 | -1 | -1 | -1 | -1 |       |
| 2   | 1  | -1 | -1 | -1 | 1  |       |
| 3   | -1 | 1  | -1 | -1 | 1  |       |
| 4   | 1  | 1  | -1 | -1 | -1 |       |
| 5   | -1 | -1 | 1  | -1 | 1  |       |
| 6   | 1  | -1 | 1  | -1 | -1 |       |
| 7   | -1 | 1  | 1  | -1 | -1 |       |
| 8   | 1  | 1  | 1  | -1 | 1  |       |
| 9   | -1 | -1 | -1 | 1  | 1  |       |
| 10  | 1  | -1 | -1 | 1  | -1 |       |
| 11  | -1 | 1  | -1 | 1  | -1 |       |
| 12  | 1  | 1  | -1 | 1  | 1  |       |
| 13  | -1 | -1 | 1  | 1  | -1 |       |
| 14  | 1  | -1 | 1  | 1  | 1  |       |
| 15  | -1 | 1  | 1  | 1  | 1  |       |
| 16  | 1  | 1  | 1  | 1  | -1 |       |

- (d) Suppose that after analyzing the data, the two largest effects were  $B$  and  $E$ . Dropping the negligible factors results in what type of design in the remaining factor
- (e) Suppose that after analyzing the data, the three largest effects were  $A$ ,  $C$ , and  $E$ . Dropping the negligible factors results in what type of design in the remaining factors?
- (f) Suppose that the original experimental situation would only allow eight runs to be made on each day. You wish to consider using the two days as two blocks. What factor (and its aliases) would you choose to confound with blocks? Use the column labeled “Block” in the design matrix above to identify which runs would be assigned to block 1 and which runs would be assigned to block 2.
- (g) Suppose it was possible to conduct only 8 runs for the initial experiment. What design would you recommend? What will be the resolution of the design and what will be the alias structure?

**4.34** An unreplicated  $2^{5-1}$  fractional factorial experiment has been run in a chemical process. The response variable is viscosity. The engineer has used the following factors:

| Factor            | Natural Levels                  | Coded Levels ( $x$ 's) |
|-------------------|---------------------------------|------------------------|
| A—time            | 40, 60 (minutes)                | -1, 1                  |
| B—temperature     | 125, 175 ( $^{\circ}\text{C}$ ) | -1, 1                  |
| C—catalyst amount | 10, 20 (grams)                  | -1, 1                  |
| D—concentration   | 60, 80 (percent)                | -1, 1                  |
| E—feed rate       | 8, 12 (g/min)                   | -1, 1                  |

After conducting the experiment, the engineer finds that three of the main effects and one of the interactions are large enough to be considered significant. The effect estimates are  $A = 60$ ,  $B = 56$ ,  $C = 40$ , and  $AC = 32$ . The overall average of the 16 runs in this experiment is 1500.

- (a) Write down a model that could be used for predicting viscosity over the region of exploration from this experiment. Write the model in the coded variables.
- (b) Write the model in part (a) above in terms of the natural variables.
- (c) Predict viscosity at the point time = 50, temperature = 165, catalyst amount = 20, concentration = 75, and feed rate = 12.
- (d) In the previous question, does it matter what value is given for feed rate?

### Short Answer Questions

**4.35** In a resolution IV design:

- (a) All main effects are aliased with three-factor interactions.
- (b) All two factor interactions are aliased with at least one other two-factor interaction.
- (c) Main effects and two-factor interactions are clear of each other.
- (d) All of the above statements are true.
- (e) None of the above statements are true.



- 4.36** In the  $2^{7-3}$  fractional factorial design with  $E = ABC$ ,  $F = BCD$ , and  $G = ACD$ , if factors  $A$ ,  $B$ , and  $C$  are dropped, the remaining factors form the following:
- (a) An unreplicated full factorial design.
  - (b) A fractional factorial design in four factors.
  - (c) A replicated full factorial design in 16 runs.
  - (d) None of the above statements are true.
- 4.37** In a resolution III design:
- (a) At least one main effect is aliased with some two-factor interactions.
  - (b) Some two-factor interactions may be aliased with other two-factor interactions.
  - (c) A screening design with lots of factors may be a good choice.
  - (d) All of the above statements are true.
  - (e) None of the above statements are true.
- 4.38** When there are several variables, the system or process is likely to be driven primarily by some of the main effects and low interactions. This idea is referred to as:
- (a) The blocking effect
  - (b) The projection property
  - (c) The sparsity of effects principle
  - (d) Sequential experimentation
- 4.39** Every  $2^{k-p}$  fractional factorial contains full factorials in some subsets of the original  $k$  factors. This is called
- (a) The projection property
  - (b) Sequential experimentation
  - (c) Ockham's razor
  - (d) The sparsity of effects principle
  - (e) Design resolution
- 4.40** Consider a  $2^{4-1}$  fractional factorial design. If the principal fraction is run first ( $I = ABCD$ ) and then later augmented with the alternate fraction, the four-factor interaction effect is confounded with blocks. **True False**
- 4.41** The resolution of a regular two-level fractional factorial design is the number of words in the defining relation. **True False**
- 4.42** Every fractional factorial contains full factorials in some subsets of the original  $k$  factors. This is called design projection. **True False**
- 4.43** When there are several variables, the system or process is likely to be driven primarily by some of the main effects and low-order interactions. This idea is referred to as sparsity of effects. **True False**
- 4.44** A  $2^{k-p}$  Fractional factorial design has 4 center runs. The observations are 40, 42, 38, and 4. The pure error sum of squares is
- (a) 21
  - (b) 18

- (c) 24  
(d) None of the above
- 4.45 Not counting the principal fraction, how many alternative fractions are there in a  $2^{k-p}$  as a function of  $p$ ?
- 4.46 With a resolution  $R$  design, main effects are confounded with interactions of order  $R - 1$ , two-factor interactions are confounded with interactions of order  $R - 2$ , and so forth. **True False**
- 4.47 For a half fraction of a two-level factorial design the maximum resolution possible is equal to the number of factors. **True False**
- 4.48 For a  $2^{10-2}$  design the best resolution attainable is 5. **True False**
- 4.49 Nongeometric Plackett-Burman designs are of resolution IV. **True False**
- 4.50 The design points of the  $2^{k-p}$  family are at the corners of a cube in a  $k$ -dimensional space and they project into a cuboidal region in any subset of the original  $k$  factors. **True False**

---

# 5

---

## PROCESS IMPROVEMENT WITH STEEPEST ASCENT

In previous chapters we dedicated considerable attention to constructing and applying designed experiments with a view toward model building. The designs discussed are very efficient for development of empirical equations that relate controllable factors to an important response. Clearly these empirical regression equations serve to provide information about the properties of the system from which the data are taken. Signs and magnitudes of coefficients and the presence or absence of interaction in the system underscore important pieces of information for the user. Often the primary purpose of the model is for interpretations such as these.

There are many other situations in which a model is used for **process optimization** or **process improvement**. The result of a model-building procedure is an equation. A statistical analyst is armed with mathematical and analytical techniques, the purpose of which is to optimize the process. This and other chapters that follow will deal with process optimization for a variety of situations. Here we assume that the practitioner is experimenting with a system (perhaps a new system) in which the goal is not to find a *point of optimum response*, but to search for a new *region* in which the process or product is improved. Perhaps the current working region for designed experiments is only based on an educated guess, or on preliminary experiments on a different manufacturing scale (such as a pilot plant). It is felt that improvement can be found. For example, in a chemical process one seeks to find a new region where improved yields will be experienced.

The experimental design, model-building procedure, and sequential experimentation that are used in searching for a region of improved response constitute the **method of steepest ascent**. The type of designs that are most frequently used are those discussed in Chapters 3 and 4, namely, two-level factorial and fractional factorial designs. One must keep in mind that the strategy involves sequential movement from one region in the factors to another. As a result the total operation may involve more than one experiment. Thus,

design economy and model simplicity may be very important. One begins by assuming that a first-order model (a planar representation) is a reasonable approximation of the system in the initial region of  $x_1, x_2, \dots, x_k$ . Then the method of steepest ascent consists of the following steps:

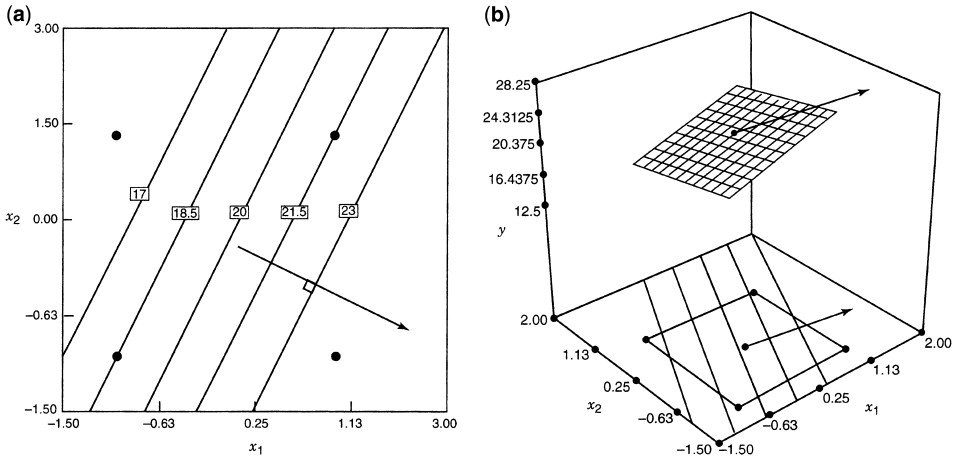
1. Fit a **first-order model** (a plane or hyperplane) using an orthogonal design. Two-level designs are quite appropriate, although center runs are often recommended.
2. Compute a path of **steepest ascent** if maximizing the response is required. If minimum response is required, one should compute the path of **steepest descent**. The path of steepest ascent is computed so that one may expect the **maximum increase** in response per distance moved in the design space. Steepest descent produces a path that results in a **maximum decrease** in response.
3. Conduct **experimental runs** along the path. That is, do either single runs or replicated runs, and observe the response value. The results will normally show improving values of the response. At some region along the path the improvement will decline and eventually disappear. This stems from the deterioration of the simple first-order model once one strays too far from the initial experimental region. A good rule of thumb is to continue experimentation along the path of steepest ascent until two consecutive runs result in decreased response values. Often the first experimental run should be taken near the design perimeter (at coordinates corresponding to a value of 1.0 in an important variable) to serve as a confirming experiment.
4. At some point where an approximation of the maximum (or minimum) response is located on the path, a **base** for a **second experiment** is chosen. The design should again be a first-order design. It is quite likely that center runs for testing curvature, as well as degrees of freedom for interaction-type lack of fit, are important at this point.
5. A second experiment is conducted, and another first-order model is fitted to the data. A test for lack of fit is made. If the lack of fit is not significant, a second path based on the new model is computed. This is often called a **mid-course correction**. Single or replicated experiments along this second path are conducted. It is quite likely that the rate of improvement will not be as strong as that enjoyed in the first path. After improvement is diminished, or some lack of fit is detected, one typically has a base for conducting a more elaborate experiment and a more sophisticated process optimization.

This step-by-step procedure is meant to be only a guideline. It is quite possible that only one stage (and hence one path) will be used. Clearly, if interaction or quadratic lack-of-fit contributions are prominent at the second stage, the analyst will likely not conduct experiments along a path that is based on a planar model. Augmentation of the first-order design to allow higher-order models will be discussed in more detail in Chapter 6.

## 5.1 DETERMINING THE PATH OF STEEPEST ASCENT

### 5.1.1 Development of the Procedure

The coordinates along the path of steepest ascent depend on the **signs** and **magnitudes** of the regression coefficients in the fitted first-order model. The following describes the



**Figure 5.1** The path of steepest ascent. (a) Contour plot. (b) Three-dimensional response surface view.

movement in, say,  $x_j$  ( $j = 1, 2, \dots, k$ ) relative to the movement of the other factors. Remember that the variables are in coded form with the center of the design at  $x_1 = x_2 = \dots = x_k = 0$ .

The movement in  $x_j$  along the path of steepest ascent is proportional to the magnitude of the regression coefficient  $b_j$  with the direction taken matching the sign of the coefficient. Steepest descent requires the direction to be opposite the sign of the coefficient.

For example, if the experiment and the resulting first-order model analysis produces an equation  $\hat{y} = 20 + 3x_1 - 1.5x_2$ , the path of the steepest ascent will result in  $x_1$  moving in a positive direction and  $x_2$  in a negative direction. In addition,  $x_1$  will move twice as fast; that is,  $x_1$  moves two units for every single unit movement in  $x_2$ . Figure 5.1 indicates the nature of the path of steepest ascent for this example. The path is indicated by the arrow. Note that the path is **perpendicular** to lines of constant response. This is clearly the most rapid movement (steepest ascent) toward large response values. For  $k \geq 3$ , these lines become planes (or hyperplanes) and the path of steepest ascent moves perpendicular to these planes.

While a path created by moving  $x_j$  a relative distance that is proportional to the regression coefficient  $b_j$  is a reasonable and intuitive selection, one may appreciate and better understand the procedure through a mathematical development of the procedure. Consider the fitted first-order regression model

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k$$

By the path of **steepest ascent** we mean that which produces a maximum estimated response with the constraint that  $\sum_{i=1}^k x_i^2 = r^2$ . Constrained optimization must be used, because maximizing the fitted response  $\hat{y}$  from the first-order model results in all of the  $x$ 's at infinity. In other words, of all points that are a fixed distance  $r$  from the center of the design, we seek that point  $x_1, x_2, x_3, \dots, x_k$  for which  $\hat{y}$  is maximized. Remember that in the metric of the coded design variables the design center is  $(0, 0, \dots, 0)$ . As a result, the constraint given by  $\sum_{i=1}^k x_i^2 = r^2$  is that of a sphere with radius  $r$ .

The solution to this optimization problem involves the use of Lagrange multipliers. Maximization requires the partial derivatives with respect to  $x_j$  ( $j = 1, 2, \dots, k$ ) of

$$L = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k - \lambda \left( \sum_{i=1}^k x_i^2 - r^2 \right) \tag{5.1}$$

The derivative with respect to  $x_j$  is

$$\frac{\partial L}{\partial x_j} = b_j - 2\lambda x_j \quad (j = 1, 2, \dots, k)$$

Setting  $\partial L/\partial x_j = 0$  gives the following coordinate of  $x_j$  of the path of steepest ascent

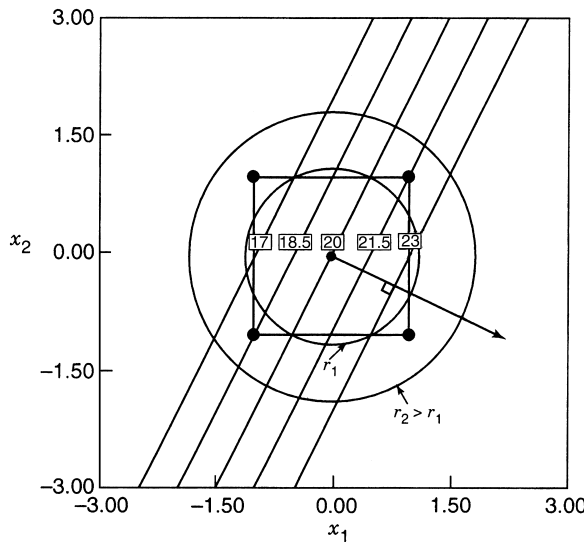
$$x_j = \frac{b_j}{2\lambda} \quad (j = 1, 2, \dots, k)$$

Now, the quantity  $1/2\lambda = \rho$  may be viewed as a **constant of proportionality**. That is, the coordinates are given by

$$x_1 = \rho b_1, \quad x_2 = \rho b_2, \quad \dots, \quad x_k = \rho b_k \tag{5.2}$$

where, for **steepest ascent**, the constant  $\rho$  is **positive**. For **steepest descent**,  $\rho$  is taken to be **negative**. Now, this implies that the choice of  $\rho$ , which is related to  $\lambda$ , merely determines the distance from the design center that the resulting point will reside. As a result, of course, the constant  $\rho$  is determined by the practitioner.

One can view the path of steepest ascent as being generated as shown in Fig. 5.2. Suppose again that  $\hat{y} = 20 + 3x_1 - 1.5x_2$ . The points on the path are viewed as locations of



**Figure 5.2** The path of steepest ascent as the path passing through the maximum value of  $y$  at a fixed distance  $r$  from the design center.

maximum values of  $\hat{y}$  at fixed radii. Each point on the path then, is a point of maximum response for a given sized step from the center of the current design. Again, note that the path involves a movement in a positive direction for  $x_1$  and a negative direction for  $x_2$ , with a change of two units in  $x_1$  for every single unit change in  $x_2$ .

### 5.1.2 Practical Application of the Method of Steepest Ascent

We now give two examples illustrating the calculations to determine the path of steepest ascent. We also give some practical guidance regarding its use.

**Example 5.1 The Plasma Etch Process** A common processing step in semiconductor manufacture is plasma etching of silicon wafers. The etch rate is the typical response of interest. Table 5.1 presents data from a  $2^2$  factorial design with four center points used to study this process. The process variables of interest are the anode–cathode gap ( $x_1$ ) and the power applied to the cathode ( $x_2$ ). The etch rates observed here are too low, and the experimenter would like to move to a region where the etch rate is around 1000 Å/min.

Table 5.2 shows the condensed analysis of this experiment from Design-Expert. Notice that the interaction term is not significant and there is no indication of lack of fit. Therefore, we will use the first-order model

$$\hat{y} = 766.25 - 66.25x_1 + 43.75x_2$$

to construct the path of steepest ascent.

Since the sign of  $x_1$  is negative and the sign of  $x_2$  is positive, we will decrease the gap and increase the power in order to increase the etch rate. Furthermore, for every unit of change in the gap ( $x_1$ ), we will change the power by  $43.75/66.25 = 0.66$  units. Consequently, if we choose the gap step size *in coded units* to be  $\Delta x_1 = -1.0$ , then the power step size *in coded units* is  $\Delta x_2 = 0.66$ . In *natural units* the step sizes are  $\Delta \text{gap} = -0.20$  cm and  $\Delta \text{power} = 16.5$  W.

Table 5.3 and Fig. 5.3 show the results of applying steepest ascent to this process. The first point on the path of steepest ascent, denoted Base + $\Delta$  in Table 5.3, has the setting gap = 1.20 cm and power = 316.5 W. The observed response at this point is  $y = 845$  Å/min. Notice that this is very close to the 840-Å/min contour passing near to this point. It is always

**TABLE 5.1 The Plasma Etch Experiment**

| Gap (cm) | Power (W) | $x_1$ | $x_2$ | $y$ (Etch Rate Å/min) |
|----------|-----------|-------|-------|-----------------------|
| 1.20     | 275       | -1    | -1    | 775                   |
| 1.60     | 275       | +1    | -1    | 670                   |
| 1.20     | 325       | -1    | +1    | 890                   |
| 1.60     | 325       | +1    | +1    | 730                   |
| 1.40     | 300       | 0     | 0     | 745                   |
| 1.40     | 300       | 0     | 0     | 760                   |
| 1.40     | 300       | 0     | 0     | 780                   |
| 1.40     | 300       | 0     | 0     | 720                   |

**TABLE 5.2 Analysis of the Etch Rate Experiment**

Response: Etch rate

ANOVA for Selected Factorial Model  
Analysis of Variance Table [Partial sum of squares]

| Source     | Sum of Squares | DF             | Mean     | F-Value | Prob > F |  |
|------------|----------------|----------------|----------|---------|----------|--|
| Model      | 25968.75       | 3              | 8656.25  | 13.53   | 0.0300   |  |
| A          | 17556.25       | 1              | 17556.25 | 27.45   | 0.0135   |  |
| B          | 7656.25        | 1              | 7656.25  | 11.97   | 0.0406   |  |
| AB         | 756.25         | 1              | 756.25   | 1.18    | 0.3564   |  |
| Curvature  | 450.00         | 1              | 450.00   | 0.70    | 0.4632   |  |
| Pure Error | 1918.75        | 3              | 639.58   |         |          |  |
| Cor Total  | 28337.50       | 7              |          |         |          |  |
| Std. Dev.  | 25.29          | R-Squared      |          | 0.9312  |          |  |
| Mean       | 758.75         | Adj R-Squared  |          | 0.8624  |          |  |
| C.V.       | 3.33           | Adeq Precision |          | 11.004  |          |  |

| Factor       | Coefficient Estimate | DF | Standard Error | 95% CI Low | 95% CI High | VIF  |
|--------------|----------------------|----|----------------|------------|-------------|------|
| Intercept    | 766.25               | 1  | 12.64          | 726.01     | 806.49      |      |
| A-gap        | -66.25               | 1  | 12.64          | -106.49    | -26.01      | 1.00 |
| B-Power      | 43.75                | 1  | 12.64          | 3.51       | 83.99       | 1.00 |
| AB           | -13.75               | 1  | 12.64          | -53.99     | 26.49       | 1.00 |
| Center Point | -15.00               | 1  | 17.88          | -71.91     | 41.91       | 1.00 |

Final Equation in Terms of Coded Factors:

$$\begin{aligned} \text{Etch rate} = & +766.25 \\ & -66.25 * A \\ & +43.75 * B \\ & -13.75 * A * B \end{aligned}$$

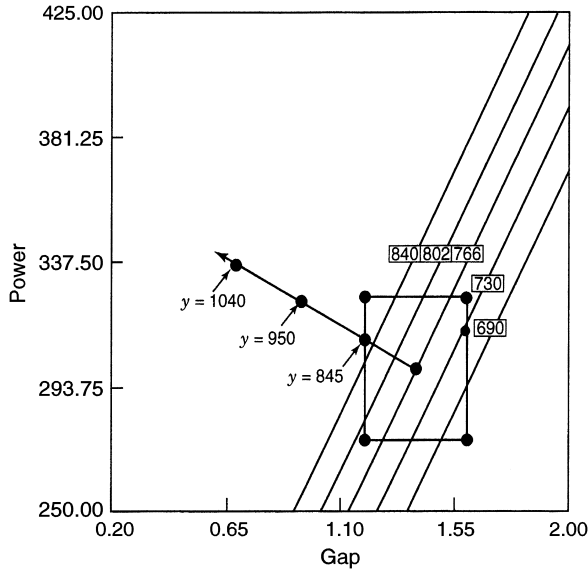
Final Equation in Terms of Actual Factors:

$$\begin{aligned} \text{Etch rate} = & -450.00000 \\ & +493.75000 * \text{gap} \\ & +5.60000 * \text{power} \\ & -2.75000 * \text{gap} * \text{power} \end{aligned}$$

**TABLE 5.3 Computation of the Path of Steepest Ascent for the Plasma Etch Process, Example 5.1**

| Point                 | Coded Variables |           | Natural Variables |       |      |
|-----------------------|-----------------|-----------|-------------------|-------|------|
|                       | Gap (cm)        | Power (W) | $x_1$             | $x_2$ | y    |
| Base (starting point) | 1.40            | 300       | 0                 | 0     |      |
| $\Delta$              | -0.20           | 16.5      | -1                | 0.66  |      |
| Base + $\Delta$       | 1.20            | 316.5     | -1                | 0.66  | 845  |
| Base + 2 $\Delta$     | 1.00            | 333       | -2                | 1.32  | 950  |
| Base + 3 $\Delta$     | 0.80            | 349.5     | -3                | 1.98  | 1040 |





**Figure 5.3** The path of steepest ascent for the plasma etch experiment in Example 5.1.

a good idea to run the first point on the path near to the original experimental region as a **confirmation test**, to ensure that conditions experienced during the original experiment have not changed. At Base +2Δ the etch rate response has increased to 950 Å/min, and at base +3Δ it has reached 1040 Å/min. Clearly we have arrived at a region close to the desired etch rate of 1000 Å/min.

Because the plasma etch experiment in Example 5.1 has only two process variables, an experimenter could actually implement the method of steepest ascent **graphically**, just by reference to a contour plot such as Fig. 5.3. However, when there are more than  $k = 2$  design factors, a formal procedure can be helpful.

It is easy to give a general algorithm for determining the coordinates of a point on the path of steepest ascent. Assume that the point  $x_1 = x_2 = \dots = x_k = 0$  is the base or origin point. Then

1. Choose a step size in one of the process variables, say  $\Delta x_i$ . Usually, we select the variable we know the most about, or we select the variable that has the largest (or nearly the largest) absolute regression coefficient  $|b_i|$ .
2. Calculate the step size in the other variables with (from Eq. 5.2)

$$\Delta x_j = \frac{b_j}{b_i/\Delta x_i}, \quad j = 1, 2, \dots, k, \quad i \neq j \tag{5.3}$$

3. Convert the  $\Delta x_j$  from the coded variables to the natural variables.

We will illustrate this procedure with a four-variable example.

**Example 5.2 A Four-Variable Illustration of Steepest Descent** Many manufacturing companies use injection-molded parts as components of their product. Shrinkage is often a problem in injection molding. Often a die for a part will be built larger than nominal to allow for part shrinkage. In the following experiment a new part is being produced, and ultimately it is important to find the proper settings to minimize shrinkage. In the following experiment, the response values are deviations from nominal—that is, shrinkage. The factor levels in natural and design units are as follows:

| Factor                                   | Design Units |      |
|--|--------------|------|
|  | -1           | +1   |
| $x_1$ : Injection velocity (ft/sec)      | 1.0          | 2.0  |
| $x_2$ : Mold temperature ( $^{\circ}$ C) | 100          | 150  |
| $x_3$ : Mold pressure (psi)              | 500          | 1000 |
| $x_4$ : Back pressure (psi)              | 75           | 120  |

The design was an unreplicated  $2^4$  factorial. The response is in units of  $10^{-4}$  cm. The actual data are not given, but the first-order model fit to the data is

$$\hat{y} = 80 - 5.28x_1 - 6.22x_2 - 1.21x_3 - 1.07x_4$$

This linear regression model relates predicted shrinkage to the design variables in coded design units. It appears that movement to a region with higher injection velocity and mold temperature will be beneficial. In addition, a slight increase in mold pressure and back pressure may be beneficial. One must keep in mind that we are searching for the path of steepest descent. Then an increase proportional to the regression coefficient in each factor will define the proper path for future experiments.

Suppose that we select  $x_1$  as the variable to define the step size (note that  $x_1$  has one of the two largest regression coefficients). We will let  $\Delta x_1 = 1$ , corresponding to 0.5 ft/sec in injection velocity. Then from Equation 5.3 we have

$$\Delta x_2 = \frac{b_2}{b_1} = \frac{6.22}{5.28} = 1.178 \text{ design units increase in } x_2$$

$$\Delta x_3 = \frac{b_3}{b_1} = \frac{1.21}{5.28} = 0.23 \text{ design units increase in } x_3$$

$$\Delta x_4 = \frac{b_4}{b_1} = \frac{1.07}{5.28} = 0.203 \text{ design units increase in } x_4$$

The corresponding change in terms of the natural units are

$$\Delta(\text{mold temperature}) = (1.178)[(150 - 100)/2] = (1.178)(25) = 29.45$$

$$\Delta(\text{mold pressure}) = (0.23)(250) = 57.5$$

$$\Delta(\text{back pressure}) = (0.203)(22.5) = 4.57$$

**TABLE 5.4 The Path of Steepest Descent for Example 5.2**

|                      | Coded Units |       |       |       | Natural Units |        |       |        |
|----------------------|-------------|-------|-------|-------|---------------|--------|-------|--------|
|                      | $x_1$       | $x_2$ | $x_3$ | $x_4$ | ft/sec        | °C     | psi   | psi    |
| Base                 | 0           | 0     | 0     | 0     | 1.5           | 125    | 750   | 97.50  |
| Increment = $\Delta$ | 1.0         | 1.178 | 0.23  | 0.203 | 0.5           | 29.45  | 57.5  | 4.57   |
| Base + $\Delta$      | 1.0         | 1.178 | 0.23  | 0.203 | 2.0           | 154.45 | 807.5 | 102.07 |
| Base + 2 $\Delta$    | 2.0         | 2.356 | 0.46  | 0.406 | 2.5           | 183.90 | 865.0 | 106.64 |
| Base + 3 $\Delta$    | 3.0         | 3.534 | 0.69  | 0.609 | 3.0           | 213.35 | 922.5 | 111.21 |
| Base + 4 $\Delta$    | 4.0         | 4.712 | 0.92  | 0.812 | 3.5           | 242.80 | 980.0 | 115.78 |

Table 5.4 shows the path of steepest descent in terms of both coded (design) units and natural units. The strategy involves making experimental runs along this path until no further improvement in shrinkage is observed.

### *Practical Notes Regarding Steepest Ascent*

1. Steepest ascent is a first-order **gradient-based** optimization technique. It works very well when starting a long way from the optimum. When used near an extreme point on the true response surface, steepest ascent will usually result in a very short movement away from the starting point. This could be an indication to consider expanding the model by adding higher-order terms.
2. We have illustrated making **individual observations** at each point on the path. In some cases **replicates**, or **repeat runs**, will be useful. For example, if we are optimizing a process such as chemical vapor deposition, the variability in layer thickness will generally increase with the average thickness of the deposited layer. Using repeat measurements at different sites on the unit or processing several units at the same time and determining both the average and the standard deviation of the response can be very useful.
3. Other first-order optimization techniques can be employed as alternatives to steepest descent. These more automatic **hill-climbing procedures** do not provide the feedback of process information about factor effects obtained from the factorial design.
4. Some experimenters will adjust the step size after the first few steps, depending on the results obtained. For example, if the response is changing slowly, the step size can be lengthened. This should be done very carefully, however, as it is easy to overshoot the region of the optimum if the step size is too big. If no other information is available, choosing the variable with the largest coefficient for setting  $\Delta x$  may be an appropriate conservative choice.

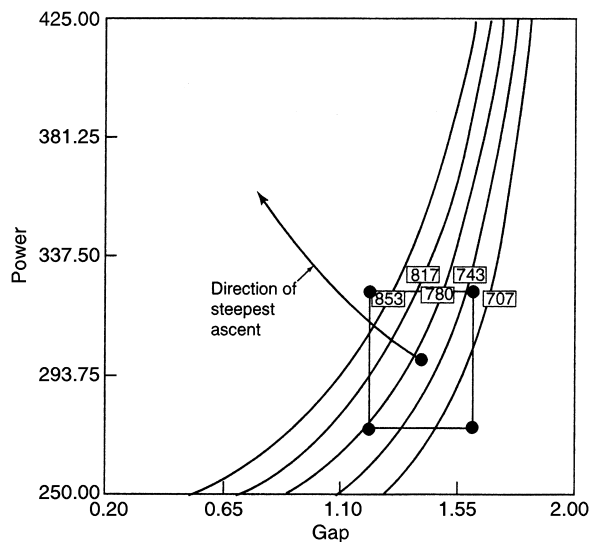
## **5.2 CONSIDERATION OF INTERACTION AND CURVATURE**

As we indicated earlier in this chapter, typically no more than two rounds of the steepest ascent (or descent) procedure will be needed. This general strategy is at its best when the

researcher begins experimentation far from the region of optimum conditions. Here one expects the first-order approximation to be quite reasonable. As the experimental region moves near the region of optimum conditions, it is expected that curvature will be more prevalent and, of course, interactions among the factors will become more important. As a result, the experimental design used to carry out the strategy should allow estimation (and testing) of potentially important interactions. In later stages of the strategy, the design should allow some information regarding model curvature. The use of center runs allows a single-degree-of-freedom estimate of quadratic curvature (see Section 3.6). Clearly, if interactions are found to be important and/or a test for curvature finds significant quadratic terms, the researcher will suspect that the steepest ascent (descent) methodology will become ineffective. Augmentation of the design to allow fitting a complete second-order model should then be done.

When second-order terms describing interaction and pure quadratic curvature ( $x_1^2, x_2^2, \dots$ ) begin to **dominate**, then continuing the ascent exercise and experimentation will be self-defeating. However, the question arises, "What do we mean by 'dominant'?" It is possible that second-order terms can be statistically significant, yet the first-order approximation allows a reasonably successful experimental strategy. One must keep in mind that "statistical significance" only implies that the effects are real in comparison with experimental error. The second-order effects may be small in magnitude compared to their first-order counterparts. As a result, there will certainly be situations where one should compute the path and take experimental trials even though certain second-order effects are significant.

Figure 5.4 shows the contour plot for the plasma etch experiment in Example 5.1 with the interaction term included in the model. The path of steepest ascent is now the curved path shown in the figure. Here the interaction is quite small, so the effect of ignoring it is negligible. Even when interactions are moderately large, ignoring them will usually cause



**Figure 5.4** The plasma etch experiment in Example 5.1 with the interaction term included in the model.

the computed path of steepest ascent to differ only modestly from the true path. In practice, this should have little effect on the final outcome. It can lead to an extra round of steepest ascent to compensate for the errors in estimating the path.

**Example 5.3 Another Four-Variable Illustration of Steepest Ascent** For this example, we consider an illustration of a steepest ascent experiment in which it is of interest to maximize the reaction yield. Four factors, *A* (amount of reactant A), *B* (reaction time), *C* (amount of reactant C), and *D* (temperature), are being considered. The natural and coded levels are given as follows:

|                | Coded Levels |    |                     |
|----------------|--------------|----|---------------------|
|                | -1           | +1 |                     |
| Natural Levels | 10           | 15 | (factor A, grams)   |
|                | 1            | 2  | (factor B, minutes) |
|                | 25           | 35 | (factor C, grams)   |
|                | 75           | 85 | (factor D, °C)      |

A  $2^{4-1}$  fractional factorial was used as the design, with the yield values given as follows:

|             |      |               |      |
|-------------|------|---------------|------|
| (1):        | 62.0 | <i>ad</i> :   | 61.8 |
| <i>ab</i> : | 69.0 | <i>bc</i> :   | 64.7 |
| <i>cd</i> : | 57.0 | <i>bd</i> :   | 62.2 |
| <i>ac</i> : | 64.5 | <i>abcd</i> : | 66.3 |

The fitted linear regression model is given by

$$\hat{y} = 63.44 + 1.9625x_1 + 2.1125x_2 - 0.3125x_3 - 1.6125x_4$$

The basis for computation of points along the path was chosen to be 1 gram of reactant A. This corresponds to  $1/2.5 = 0.4$  design units. As a result, the corresponding movements in the other design variables are  $(2.1125/1.9625)(0.4) = 0.4306$  design units for  $x_2$ ,  $(-0.3125/1.9625)(0.4) = -0.0637$  design units for  $x_3$ , and  $(-1.6125/1.9625)(0.4) = -0.3287$  design units for  $x_4$ . The movement along the path will be positive for  $x_1$  and  $x_2$ , and negative for  $x_3$  and  $x_4$ . Table 5.5 shows the appropriate coordinates along the path in the natural variables. At some location along the design perimeter, experimental runs should be made. In this case, runs were made at base  $+4\Delta$  (after four increments): Runs 9, 10, 11, and 12 indicate new experimental runs. Theoretically, one expects an increase in response as runs are taken along the path. Eventually, of course, deterioration should occur when the first-order approximation that produced the original path is no longer valid because the true response surface is now influenced by interaction and quadratic curvature. In this case a reduction in yield is experienced after run 11. Any further follow-up experiment should involve a designed experiment centered in the vicinity of run 11.

**TABLE 5.5** Coordinates on the Path of Steepest Ascent in Natural Variables for Example 5.3

| Run                  | $x_1$ | $x_2$                    | $x_3$                    | $x_4$                    | $y$  |
|----------------------|-------|--------------------------|--------------------------|--------------------------|------|
| Base                 | 12.5  | 1.5                      | 30                       | 80                       |      |
| $\Delta$             | 1.0   | (0.4306)(0.5)<br>= 0.215 | (-0.0637)(5)<br>= -0.319 | (-0.3287)(5)<br>= -1.643 |      |
| Base + $\Delta$      | 13.5  | 1.715                    | 29.681                   | 78.357                   |      |
| Base + 2 $\Delta$    | 14.5  | 1.930                    | 29.362                   | 76.714                   |      |
| Base + 3 $\Delta$    | 15.5  | 2.145                    | 29.043                   | 75.071                   |      |
| 9 Base + 4 $\Delta$  | 16.5  | 2.360                    | 28.724                   | 73.428                   | 74.0 |
| 10 Base + 6 $\Delta$ | 18.5  | 2.790                    | 28.086                   | 70.142                   | 77.0 |
| 11 Base + 8 $\Delta$ | 20.5  | 3.220                    | 27.448                   | 73.856                   | 81.0 |
| 12 Base + 9 $\Delta$ | 21.5  | 3.435                    | 27.129                   | 65.213                   | 78.7 |

**5.2.1 What About a Second Phase?**

As we indicated earlier, one should expect that the evidence of curvature in the system and interaction among the factors will eventually necessitate an abandonment of steepest ascent. Any additional phases (or mid-course corrections) of the procedure beyond the first will usually not bring about the level of success enjoyed in the initial phase. In addition, one should be careful to use a design in the second phase that allows for testing lack of fit that includes interaction and curvature induced by quadratic terms. In Example 5.3 a reasonable design for the second stage is the other half of the original  $2^{4-1}$  design (forming a complete  $2^4$ ) augmented by **center runs**. This allows for three degrees of freedom for testing interaction type lack of fit and one degrees of freedom for testing curvature. If the lack of fit is significant, one might expect little or no further success with steepest ascent.

**5.2.2 What Happens Following Steepest Ascent?**

It should be emphasized at this point that quality improvement through analysis of designed experiments, when successful, is usually an **iterative** process. This is illustrated quite well in dealing with the strategy of steepest ascent. In our example the investigator may well invest in a  $2^4$  factorial with, say, five center runs for a second phase of steepest ascent. However, if curvature and interaction are found to be quite evident, the steepest ascent procedure will certainly soon be truncated. At this point the investigators will surely be interested in finding optimum conditions through the use of a fitted **second-order model**. This allows computation of estimated optimum conditions. The first-order two-level design with center runs is nicely augmented to allow estimation of second-order terms. For example, a  $2^{4-1}$  fractional factorial (resolution IV) with five center runs is augmented with the alternate fraction and axial runs as shown in Table 5.6.

The complete design allows for efficient estimation of the terms in the model

$$y = \beta_0 + \sum_{i=1}^4 \beta_i x_i + \sum_{i=1}^4 \beta_{ii} x_i^2 + \sum_{i < j=2}^4 \sum \beta_{ij} x_i x_j + \epsilon$$

This is the **second-order response surface model** and it is widely used for process optimization. The design is called the **central composite design**. Process optimization with

**TABLE 5.6 The Central Composite Design Resulting from Augmenting an Initial  $2^{4-1}$  Design**

| $x_1$ | $x_2$ | $x_3$ | $x_4$ |                |
|-------|-------|-------|-------|----------------|
| -1    | -1    | -1    | -1    | Initial design |
| 1     | 1     | -1    | -1    |                |
| 1     | -1    | 1     | -1    |                |
| 1     | -1    | -1    | 1     |                |
| -1    | 1     | 1     | -1    |                |
| -1    | 1     | -1    | 1     |                |
| -1    | -1    | 1     | 1     |                |
| 1     | 1     | 1     | 1     |                |
| 0     | 0     | 0     | 0     |                |
| 0     | 0     | 0     | 0     |                |
| 0     | 0     | 0     | 0     |                |
| 0     | 0     | 0     | 0     |                |
| 0     | 0     | 0     | 0     |                |
| 1     | -1    | -1    | -1    | Augmentation   |
| -1    | 1     | -1    | -1    |                |
| -1    | -1    | 1     | -1    |                |
| -1    | -1    | -1    | 1     |                |
| 1     | 1     | 1     | -1    |                |
| 1     | 1     | -1    | 1     |                |
| -1    | 1     | 1     | 1     |                |
| 1     | -1    | 1     | 1     |                |
| -2    | 0     | 0     | 0     |                |
| 2     | 0     | 0     | 0     |                |
| 0     | -2    | 0     | 0     |                |
| 0     | 2     | 0     | 0     |                |
| 0     | 0     | -2    | 0     |                |
| 0     | 0     | 2     | 0     |                |
| 0     | 0     | 0     | -2    |                |
| 0     | 0     | 0     | 2     |                |

the second-order model is discussed in Chapter 6. Designs for fitting second-order models, including the central composite design, are discussed in Chapters 8 and 9.

The point made here is that the total strategy of product improvement or optimization can very well involve both steepest ascent (region seeking) and more formal response surface optimization in an iterative procedure. The transition from one to the other can be made quite easily without any waste of experimental effort.

**5.3 EFFECT OF SCALE (CHOOSING RANGE OF FACTORS)**

The methodology of proceeding along the path of steepest ascent is generally a precursor to a more elaborate experimental effort and optimization involving a more sophisticated model and analysis. In Chapters 3 and 4, considerable attention was paid to two-level designs, with variable screening being an important goal. It was suggested at that point that choosing ranges on the factors is a vital decision and something that the researcher should

not take lightly. Clearly, variable screening and steepest ascent are early steps in the process optimization experience. Sloppy decisions with little forethought in these early stages may lead to very inefficient process optimization at a later stage. It should be clear by now to the reader that there is a connection between selection of ranges of the variables and the choice of scale. For example, the coding in Example 5.2 suggests a decision in which one design unit in injection velocity is 0.5 ft/sec. An **equivalent unit** in mold temperature is 25°C. This choice of scale is a decision that is made by the practitioner. Note that the regression coefficient on back pressure ( $x_4$ ) is considerably smaller in magnitude than those on  $x_1$  and  $x_2$ . Perhaps the implication is that the choice of range (and hence, choice of scale) on back pressure was incorrect. Choosing design factor ranges certainly should improve with increased experience with the system. If variable screening has already been accomplished, then one would expect a more educated choice of scale for the hill-climbing exercise of steepest ascent. One can only use the latest information available.

A change of scale does not change the direction that a factor should move along the path of steepest ascent. However, it changes the relative magnitude of movement of the factor. Suppose now we have an ideal situation with time ( $x_1$ ) and temperature ( $x_2$ ) in which the true regression structure involving yield  $y$  is as follows:

$$E(y) = \beta_0 + \beta x_1 + \beta x_2$$

where  $\beta$  is a coefficient that corresponds to a (+1, -1) scaling for ranges of temperature of 50°C and time of 1.0 hr. Suppose researcher A chooses the above ranges with (+1, -1) scaling while researcher B chooses a 50°C range of temperature, but a 0.5 range of time, again with (+1, -1) scaling on the true factors. The model that is relevant to researcher B is given by

$$E(y) = \beta_0 + \beta x_1 + (\beta/2)x_2$$

Thus the expected value of the time regression coefficient  $b_2$  for researcher B is one-half that of the same coefficient for researcher A. As a result, the relative movement in  $x_2$  with respect to  $x_1$  along the path, in design units, will be half as great. As an example, suppose design levels are as follows for the researchers:

| Researcher A |                |       |                    |    |
|--------------|----------------|-------|--------------------|----|
|              | Natural Levels |       | Coded Design Units |    |
| Temperature  | 200°F          | 250°F | -1                 | +1 |
| Time         | 1.0            | 2.0   | -1                 | +1 |

| Researcher B |                |       |                    |    |
|--------------|----------------|-------|--------------------|----|
|              | Natural Levels |       | Coded Design Units |    |
| Temperature  | 200°F          | 250°F | -1                 | +1 |
| Time         | 1.25           | 1.75  | -1                 | +1 |



Suppose both researchers use  $2^2$  factorial experimental plans. Let us assume that the steepest ascent coordinates are to be based on a change of 25°F in temperature. The steepest ascent picture is as follows:

|                  | A           |      | B           |       |
|------------------|-------------|------|-------------|-------|
|                  | Temperature | Time | Temperature | Time  |
| Base             | 225         | 1.5  | 225         | 1.5   |
| $\Delta$         | 25          | 0.5  | 25          | 0.125 |
| Base + $\Delta$  | 250         | 2.0  | 250         | 1.625 |
| Base + $2\Delta$ | 275         | 2.5  | 275         | 1.750 |

Note that the actual change in time for researcher B is **one-fourth** the change incurred by researcher A. The 0.125 incremental change for researcher B results from the fact that the change per coded unit change in  $x_2$  is only one-half that experienced by researcher A on the average. In addition, the computation of the coordinates for the natural variable must take into account that the design unit described by researcher B is only one-half that described by researcher A. As a result the reader should be able to ascertain the following general rule that reflects the distinction between steepest ascent coordinates in a  $k$ -variable problem.

Suppose researcher A chooses scale factor (range)  $r_1, r_2, \dots, r_k$  and researcher B chooses  $r'_1, r'_2, \dots, r'_k$ , where  $a_j = r_j/r'_j$ . Refer to the relative movements along the path in the natural variables as follows:  $\Delta_1, \Delta_2, \dots, \Delta_k$  and  $\Delta'_1, \Delta'_2, \dots, \Delta'_k$  for researcher A and B, respectively; then  $\Delta_j/\Delta'_j = a_j^2$  for  $j = 1, 2, \dots, k$ .

This does suggest that the user of steepest ascent must use whatever knowledge of the system is at his or her disposal to determine the range of the variable and hence the scale—that is, the definition of a design unit. As we indicated earlier, each experimental experience with a particular system allows for a more educated choice of intervals on the variable being studied. One can view this general procedure as one in which there are truly many paths of steepest ascent (or descent) that will lead to a region where the response is improved. The method itself can be a learning device that will allow for a better choice of ranges in future experiments, particularly those experiments in which the goal is to find optimum process conditions. One should not let that difficulty in choosing ranges prohibit the use of this region-seeking method. The methodology allows the user to be the beneficiary of more appropriate regions and ranges for future experiments.

**5.4 CONFIDENCE REGION FOR DIRECTION OF STEEPEST ASCENT**

It is useful to take into account sampling variation in assessing the nature of the path of steepest ascent. One must remember that the path is based on the regression coefficients

and that these coefficients have sampling properties characterized by standard errors. As a result, the path has sampling variation. This sampling variation can lead to a confidence region for the path itself. The value of the confidence region may be derived by plots, say in the case of two or three variables. A graphical analysis may indicate the amount of flexibility the practitioner has in experiments along the path. A tighter region gives the user confidence that the path is being estimated well.

Suppose there are  $k$  design variables and, indeed, coefficients  $b_1, b_2, \dots, b_k$  provide estimates of the relative movement of variables along the path. Assuming that the first-order model is correct, the **true path** is defined by parameters  $\beta_1, \beta_2, \dots, \beta_k$ , and

$$E(b_i) = \beta_i \quad (i = 1, 2, \dots, k)$$

and the true coefficients are proportional to the relative movement along the path (i.e., Eq. 5.2) implies

$$\beta_i = \gamma X_i \quad (i = 1, 2, \dots, k) \quad (5.4)$$

where  $X_i$  are the **direction cosines** of the path. In other words,  $X_1, X_2, \dots, X_k$  are constants that, if known, could be used to compute any coordinates on the true path. We can view the relationship in Equation 5.4 as a regression model *without an intercept*. The subject of the statistical inference here will be the  $X_i$ . Though the regression structure may appear to be rather unorthodox, we consider the model

$$b_i = \gamma X_i + \varepsilon_i \quad (i = 1, 2, \dots, k) \quad (5.5)$$

The variance of the  $b_i$  is constant across all coefficients if one uses a standard two-level orthogonal design. Call the estimated variance  $s_b^2$ . A second variance,  $s_b^{2*}$ , is found from the error mean square of the regression of Equation 5.5. The quantity  $s_b^{2*}$  is given by

$$s_b^{2*} = \frac{\sum_{i=1}^k (b_i - \hat{\gamma} X_i)^2}{k - 1}$$

where

$$\hat{\gamma} = \frac{\sum_{i=1}^k b_i X_i}{\sum_{i=1}^k X_i^2}$$

Then, of course,  $k - 1$  is the number of error degrees of freedom for the regression. As a result,

$$\frac{s_b^{2*}}{s_b^2} \sim F_{k-1, v_b} \quad (5.6)$$

where  $v_b$  is the number of error degrees of freedom associated with the estimate  $s_b^2$  (from the steepest ascent experiment). Values of  $X_1, X_2, \dots, X_k$  that fall inside the confidence region are those for which the resulting values of  $s_b^{2*}/s_b^2$  do not appear to refute Equation 5.6. In particular, coordinates outside the confidence region are those for which  $s_b^{2*}$  is significantly larger than  $s_b^2$ . As a result, the  $100(1 - \alpha)\%$  confidence region is defined as the set of values  $X_1, X_2, \dots, X_k$  for which

$$\sum_{i=1}^k \frac{(b_i - \hat{\gamma}X_i)^2/(k - 1)}{s_b^2} \leq F_{\alpha, k-1, v_b} \tag{5.7}$$

where  $F_{\alpha, k-1, v_b}$  is the upper  $100(1 - \alpha)\%$  point of the  $F_{k-1, v_b}$  distribution.

**What Does the Confidence Region Mean?** One must understand that specific coordinates  $X_1, X_2, \dots, X_k$  specify the direction along the path. For example,  $X_1 = \sqrt{0.5}$ ,  $X_2 = \sqrt{0.3}$ , and  $X_3 = \sqrt{0.2}$  specify a direction and also represent coordinates that are a unit distance away from the design origin. (Keep in mind that we remain in design unit scaling.) The confidence region turns out to be a cone (or a hypercone in more than three variables) with the apex at the design origin and all points a unit distance from the origin satisfying

$$\sum_{i=1}^k b_i^2 - \frac{\left(\sum_{i=1}^k b_i X_i\right)^2}{(k - 1) \sum_{i=1}^k X_i^2} \leq s_b^2 F_{\alpha, k-1, v_b} \tag{5.8}$$

**Example 5.4 An Example with  $k = 2$**  Consider the path of steepest ascent illustrated in Fig. 5.1. The coefficients are  $b_1 = 3$  and  $b_2 = -1.5$ . Suppose further that the estimates of the variances of the coefficients are both  $\frac{1}{4}$  and there are four error degrees of freedom. The value  $F_{0.05, 2, 4} = 7.71$ . As a result, the 95% confidence region for the path of steepest ascent at fixed distance  $X_1^2 + X_2^2 = 1.0$  is determined by solutions  $(X_1, X_2)$  to

$$9 + 2.25 - (3X_1 - 1.5X_2)^2 \leq \frac{1}{4}(7.71)$$

or

$$(3X_1 - 1.5X_2)^2 \geq 9.3225 \tag{5.9}$$

As a result, the total confidence region on the path is as illustrated in Fig. 5.5. Similar graphical approaches to the problem are reasonable for  $k = 3$ . See Box and Draper (1987). For  $k > 3$  one cannot display a graphical picture of the confidence region. Of course, one can always substitute any value  $(X_1, X_2, \dots, X_k)$  into Equation 5.7 to determine if the point falls inside the confidence region. In general, it may be rather difficult to determine the relative size of the confidence region, and yet the user needs to have some indication of whether or not it is permissible to continue. Box and Draper describe an interesting analytic procedure for determining what percentage of the possible directions was excluded by the 95% confidence region. This then produces some impression about how “tight” the

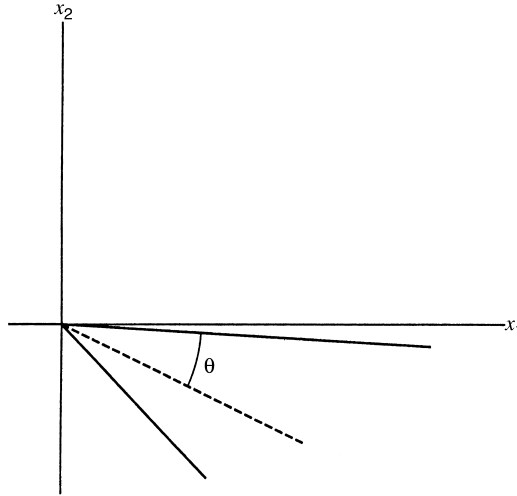


Figure 5.5 Confidence region on path of steepest ascent in Fig. 5.1.

confidence region is for any specific example. In our simple  $k = 2$  example, this might correspond to the determination of the angle  $\theta$  in Fig. 5.5. It turns out that

$$\theta = \arcsin \left\{ \frac{(k - 1)s_b^2 F_{\alpha, k-1, v_2}}{\sum_{i=1}^k b_i^2} \right\}^{1/2} \tag{5.10}$$

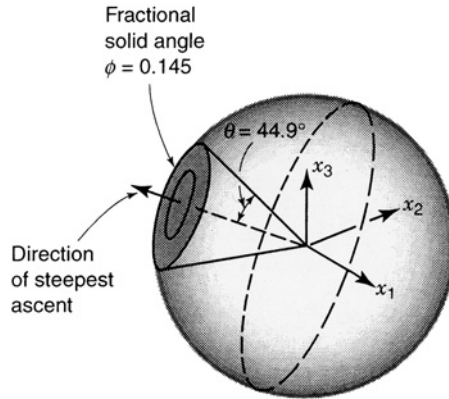
For our example, we have

$$\begin{aligned} \theta &= \arcsin \left\{ \left[ \left( \frac{1}{4} \right) (7.71) \left( \frac{1}{11.25} \right) \right]^{1/2} \right\} \\ &= \arcsin[0.413] \end{aligned}$$

Thus  $\theta \cong 24.4^\circ$ . As a result, the confidence region illustrated by Fig. 5.5 sweeps an angle of  $2\theta = 48.8^\circ$ . Of interest is the ratio  $48.8/360 = 0.135$ , suggesting that 86.5% of the possible directions taken from the design origin are excluded. Clearly, the larger the percentage excluded, the more accurate the computed path. For  $k = 3$  the confidence region is a cone and the quality of the confidence region depends on the fraction of the surface area of the total sphere that is taken up by the solid angle created by the confidence cone. Consider, for example, Fig. 5.6. The dashed line represents the computed path. The ratio of the shaded area to the total surface area of the sphere determines the quality of the computed path.

### 5.5 STEEPEST ASCENT SUBJECT TO A LINEAR CONSTRAINT

Anytime a researcher encounters a task of sequential experimentation that involves considerable movement, there is the possibility that the path will move into an area of the design



**Figure 5.6** Confidence cone for direction of steepest ascent for  $k = 3$ . [From Box and Draper (1987), with permission.]

space where one or more of the variables are not in a permissible range from an engineering or scientific point of view. For example, it is quite possible that an ingredient concentration may exceed practical limits. As a result, in many situations it becomes necessary to build the path of steepest ascent with a constraint imposed in the design variables. Suppose, in fact, that we view the constraint in the form of a boundary. That is, we are bounded by

$$c_0 + c_1x_1 + c_2x_2 + \dots + c_kx_k = 0 \tag{5.11}$$

This bound need not involve all  $k$  variables. For example,  $x_j = c_0$  may represent a boundary on a single design variable. One must keep in mind that in practice the constraint must be formulated in terms of the natural (uncoded) variables and then written in the form of coded variables for manipulation.

Figure 5.7 illustrates, for  $k = 2$ , the necessary steepest ascent procedure when a linear constraint is to be applied. The procedure is as follows:

1. Proceed along the usual path of steepest ascent until contact is made with the constraining line (or plane for  $k > 2$ ). Call this point of contact point **O**.
2. Beginning at point **O**, proceed along an adjusted or modified path.
3. Conduct experiments along the modified path, as usual, with the stopping rule based on the same general principles as those discussed in Sections 5.1 and 5.2.

**What Is the Modified Path?** Figure 5.7 clearly outlines the modified path for the  $k = 2$  illustration. In general, however, the proper direction vector is one that satisfies the constraint and still makes the greatest possible progress toward maximizing (or minimizing) the response. (Clearly the modified path in our illustration is correct.)

It turns out (and is intuitively reasonable) that the modified path is given by the direction vector

$$b_i - dc_i \quad (i = 1, 2, \dots, k) \tag{5.12}$$

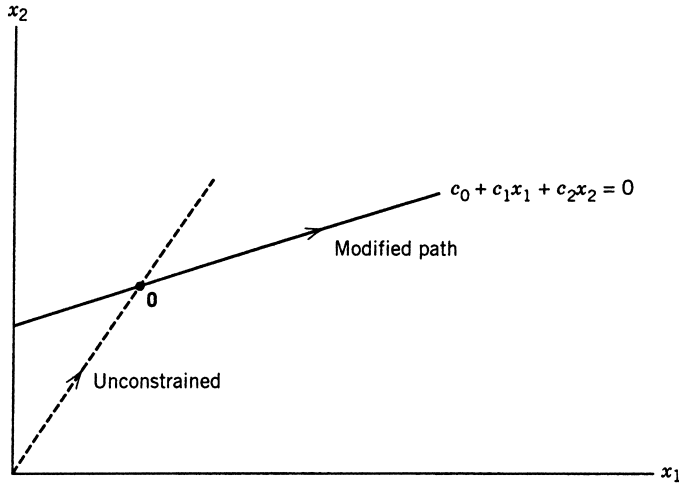


Figure 5.7 Steepest ascent with a linear constraint.

for which

$$\sum_{i=1}^k (b_i - dc_i)^2$$

is minimized. That is, the direction along the constraint line (or plane) is taken so as to be “closest” to the original path. Once again we can consider this to be a simple regression situation in which the  $b_i$  are being regressed against the  $c_i$ . As a result, the quantity  $d$  plays the role of a regression slope, and minimization of the residual sum of squares produces the value

$$d = \frac{\sum_{i=1}^k b_i c_i}{\sum_{i=1}^k c_i^2} \tag{5.13}$$

Thus, the modified path begins at point  $O$  and proceeds using the direction vector  $b_1 - dc_1, b_2 - dc_2, \dots, b_k - dc_k$ .

It remains then to determine the point  $O$ , that is, the point that lies on the original path but is also on the constrained plane. From Equation 5.2 we know that  $x_j = \rho b_j$  for  $j = 1, 2, \dots, k$ . But we also know that for point  $O$ , the constraint equation (Eq. 5.10) must hold. As a result, the collision between the original path of steepest ascent and the constraint plane must occur for  $\rho = \rho_0$  satisfying

$$c_0 + (c_1 b_1 + c_2 b_2 + \dots + c_k b_k) \rho_0 = 0$$

Thus,

$$\rho_0 = \frac{-c_0}{\sum_{i=1}^k c_i b_i} \tag{5.14}$$

As a result, the modified path starts at  $x_{j,0} = \rho_0 b_j$  (for  $j = 1, 2, \dots, k$ ) and moves along the modified path, being defined by

$$x_{j,m} = x_{j,0} + \lambda(b_j - dc_j) \quad (j = 1, 2, \dots, k) \quad (5.15)$$

where  $d$  is given by Equation 5.13 and  $\lambda$  is a proportionality constant. Note that the modified path is like the standard path except it does not start at the origin and the regression coefficient  $b_j$  is replaced by  $b_j - dc_j$  in order to accommodate the constraint.

Linear constraints on design variables occur frequently in practice. In the chemical and related fields, constraints are often imposed in situations in which the design variable represents concentrations of components in a system. In a gasoline blending system there will certainly be constraints imposed on the components in the system. General information regarding **designs for mixture problems** will be presented in Chapters 12 and 13. In the example that follows we illustrate steepest ascent in a situation where mixture factors are involved and a linear constraint must be applied to modify the path of steepest ascent.

**Example 5.5 The Fabric Strength Experiment** Consider a situation in which the breaking strength (in grams per square inch) of a certain type of fabric is a function of the amount of three important components in a kilogram of raw material. We will call the components  $\xi_1$ ,  $\xi_2$ , and  $\xi_3$ . The levels used in the experiment are as follows:

| Material | Amount in Grams |     |
|----------|-----------------|-----|
|          | -1              | +1  |
| 1        | 100             | 150 |
| 2        | 50              | 100 |
| 3        | 20              | 40  |

The remaining ingredients in the raw material are known to have no effect on fabric strength. However, it is important that  $\xi_1$  and  $\xi_2$ , the amounts of material 1 and 2 respectively, be constrained by the following equation:

$$\xi_1 + \xi_2 \leq 500$$

The design-factor centering and scaling are given by

$$\begin{aligned} x_1 &= \frac{\xi_1 - 125}{25}, & \xi_1 &= 25x_1 + 125 \\ x_2 &= \frac{\xi_2 - 75}{25}, & \xi_2 &= 25x_2 + 75 \\ x_3 &= \frac{\xi_3 - 30}{10}, & \xi_3 &= 10x_3 + 30 \end{aligned}$$

As a result, the constraint reduces to

$$25x_1 + 25x_2 \leq 300$$

Suppose the fitted regression function is given by

$$\hat{y} = 150 + 1.7x_1 + 0.8x_2 + 0.5x_3$$

Because it is desired to find the condition for increased fabric strength, the unconstrained path of steepest ascent will proceed with increasing values of all three components. The increments along the path are to correspond to changes in  $\xi_1$  of 25 g/in.<sup>2</sup>, or one design unit. This corresponds to changes in  $x_2$  and  $x_3$  of 0.47 and 0.294 units, respectively.

Now, based on the above information, the standard path of steepest ascent can be computed. But at what point does the path make contact with the constraint and with the constraint plane? From Equation 5.14 we obtain

$$\rho_0 = \frac{300}{(25)(1.7) + (25)(0.8)} = \frac{300}{62.5} = 4.8$$

As a result, the modified path starts at

$$x_{j,0} = 4.8b_j \quad (j = 1, 2, 3)$$

which results in the coordinates (8.16, 3.84, 2.4) in design units. As a result, the coordinates of the modified path are given by Equation 5.15; that is,

$$x_{j,m} = 4.8b_j + \lambda(b_j - dc_j) \quad (j = 1, 2, 3)$$

where

$$d = \frac{(25)(1.7) + (25)(0.8)}{1250} = 0.05$$

As a result, the modified path is given by

$$x_{1,m} = 8.16 + \lambda(0.45)$$

$$x_{2,m} = 3.84 + \lambda(-0.45)$$

$$x_{3,m} = 2.4 + \lambda(0.5)$$

Thus, Table 5.7 shows a set of coordinates on the path of steepest ascent, followed by points along the modified path. For the modified path we are using  $\lambda = 1.0$  for convenience. Note that all points on the modified path satisfy the constraint.

## 5.6 STEEPEST ASCENT IN A SPLIT-PLOT EXPERIMENT

In previous sections we have discussed how the method of steepest ascent operates when the experiment used to fit the first-order model is a completely randomized design (CRD). Sometimes in the steepest ascent environment we will encounter easy-to-change and hard-to-change factors. That is, the experimental situation in which the first-order model is fit



TABLE 5.7 A Constrained Path of Steepest Ascent

|                  | $x_1$    | $x_2$    | $x_3$    |
|------------------|----------|----------|----------|
| Base             | 0        | 0        | 0        |
| $\Delta$         | 1        | 0.47     | 0.294    |
| Base + $\Delta$  | 1        | 0.47     | 0.294    |
| Base + $2\Delta$ | 2        | 0.94     | 0.598    |
| $\vdots$         | $\vdots$ | $\vdots$ | $\vdots$ |
| Base + $5\Delta$ | 5        | 2.35     | 1.470    |
| Point <b>O</b>   | 8.16     | 3.84     | 2.4      |
|                  | 8.61     | 3.39     | 2.9      |
|                  | 9.06     | 2.94     | 3.4      |
|                  | 9.51     | 2.49     | 3.9      |
|                  | $\vdots$ | $\vdots$ | $\vdots$ |

is a split-plot. The split-plot nature of the experimental situation can necessitate some modification of the steepest ascent procedure. This problem is discussed in detail by Kowalski, Borror, and Montgomery (2005). Here we describe their recommendations.

We will let the hard-to-change factors be the whole-plot variables, denoted by  $z_i, i = 1, 2, \dots, k_1$  and the easy-to-change factors as the subplot variables, denoted by  $x_j, j = 1, 2, \dots, k_2$ . The fitted first-order model is

$$\hat{y} = b_0 + \sum_{i=1}^{k_1} a_i z_i + \sum_{j=1}^{k_2} b_j x_j$$

This model is used to calculate the path of steepest ascent.

Because split-plot experiments contain two types of factors, we let  $\Delta_i$  represent the step size in the whole-plot factors and  $\Delta_j$  to represent the step size in the subplot factors. There are several reasons for the choice of two distinct  $\Delta$ s:

1. It is consistent with the way the experiment was carried out; that is, in a split-plot, which separates the two types of factors.
2. The hard-to-change factors have a different standard error than the easy-to-change factors; therefore, selecting the largest regression coefficient across all factors to define a single step size could be misleading.
3. By their very nature, some factors are hard to change, and so choosing the step size for the path based on the regression coefficient of an easy-to-change factor may lead to a level that is either very difficult or not possible for one of the hard-to-change factors.

For example, it may be easier to change temperature in steps of, say,  $10^\circ\text{F}$  than to have an easy-to-change factor dictate that temperature be changed in steps of, say,  $6.42^\circ\text{F}$ . On the other hand, having a hard-to-change factor dictate the step size in the easy-to-change factors may require the steps to be larger than desired and may force the path to the boundary of the region. Although the experimenter can choose to have a common step size ( $\Delta_i = \Delta_j = \Delta$ ), we encourage the use of two separate  $\Delta$ s for the reasons mentioned above.

Let  $r_1$  represent the radius for the whole-plot factors,  $\sum_{i=1}^{k_1} z_i^2 = r_1^2$ , and let  $r_2$  represent the radius for the subplot factors,  $\sum_{j=1}^{k_2} x_j^2 = r_2^2$ . Then there are two equations involving Lagrange multipliers:

$$L_1 = b_0 + \sum_{i=1}^{k_1} a_i z_i + \lambda_1 \left( \sum_{i=1}^{k_1} z_i^2 - r_1^2 \right)$$

and

$$L_2 = b_0 + \sum_{i=1}^{k_2} b_i x_i + \lambda_2 \left( \sum_{j=1}^{k_2} x_j^2 - r_2^2 \right)$$

where  $\lambda_1$  and  $\lambda_2$  are the Lagrange multipliers. Because we are assuming a first-order model with no interactions between the whole-plot and subplot factors, these two equations can be solved separately. Therefore, taking partial derivatives of  $L_1$  with respect to each of the  $z_i$  and  $\lambda_1$  will result in the equations

$$\frac{\partial L_1}{\partial z_i} = a_i - 2\lambda_1 z_i = 0$$

and

$$\frac{\partial L_1}{\partial \lambda_1} = - \sum_{i=1}^{k_1} z_i^2 + r_1^2 = 0$$

Likewise, similar equations will result from taking the partial derivatives of  $L_2$  with respect to each  $x_j$  and  $\lambda_2$ . Hence, the solutions will be  $z_i^* = a_i/2\lambda_1$  and  $x_j^* = b_j/2\lambda_2$ . Therefore, two variables,  $z_i^*$ , and  $x_j^*$ , and their amount of change must be chosen to determine the path. There will be two centers or base points; one for the whole-plot factors (Base 1) and one for the subplot factors (Base 2). Each base point will be  $(0, 0, \dots, 0)$  in the coded variables.

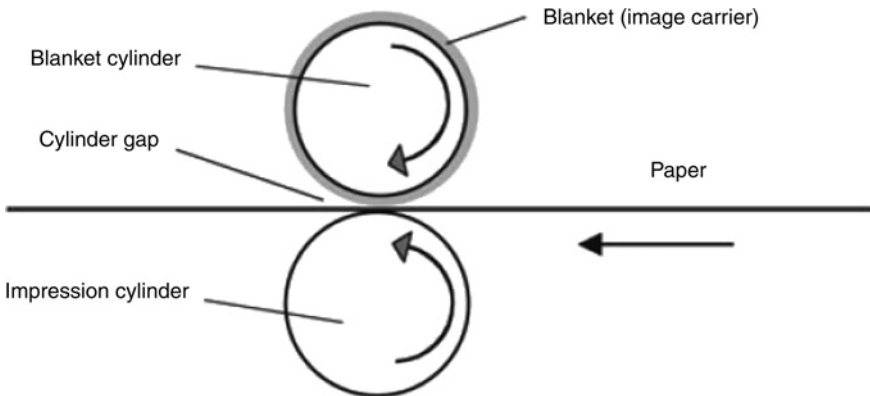


Figure 5.8 The printing press for Example 5.6.

**TABLE 5.8** Printing Press Factor Levels in the Natural Units

|      | $z_1$ | $z_2$ | $x_1$ | $x_2$ | $x_3$ |
|------|-------|-------|-------|-------|-------|
| Low  | 10    | 4     | 20    | 5     | 2     |
| High | 20    | 8     | 40    | 15    | 6     |

**Example 5.6 Image Quality** Kowalski et al. (2005) describe an experiment to study the image quality of a printing process. There are five important factors:  $A$  = blanket type (in units of thickness),  $B$  = paper type (in units of thickness),  $C$  = cylinder gap,  $D$  = ink flow, and  $E$  = press speed. Figure 5.8 is a diagram of this part of the printing press. Changing the cylinder gap, ink flow, and press speed is a very simple procedure simply consisting of making an adjustment on a control panel while the printing press is still running. Therefore, these are the easy-to-change factors in the experiment. Changing the blanket type and paper type, on the other hand, requires the press to be stopped and a manual replacement of the blanket and/or paper type. Thus, these two factors are hard-to-change factors. A completely randomized full factorial design would run the 32 treatment combinations in random order, requiring frequent stopping of the press so that the blanket type and/or paper type could be changed. Instead, a split-plot design is used with two whole plot variables,  $z_1$  and  $z_2$  ( $A$  and  $B$ ), and three subplot variables,  $x_1$ ,  $x_2$ , and  $x_3$  ( $C$ ,  $D$ , and  $E$ ). The levels of these factors in the natural units are provided in Table 5.8. Let

$$\hat{y} = 60 + 4.2z_1 + 6.8z_2 + 1.4x_1 - 3.6x_2 + 2.2x_3 \quad (5.16)$$

be the fitted first-order model. Suppose a change of two natural units in  $z_2$  (the whole-plot variable with the largest effect), equivalent to one coded unit, is chosen for determining the path in the whole-plot variables. Then  $\lambda_1 = 6.8/2(1) = 3.4$ , and the corresponding change in  $z_1$  is  $4.2/2(3.4) = 0.618$ . Now suppose that a change of five natural units in  $x_2$  (the subplot variable with the greatest effect), equivalent to one coded unit, is chosen for determining the path in the subplot variables. Then  $\lambda_2 = -3.6/2(1) = -1.8$  and the change in  $x_j$  is  $b_j/2(|\lambda_2|)$ . So the change in  $x_3$  is  $2.2/2(1.8) = 0.611$  and the change in  $x_1$  is  $1.4/2(1.8) = 0.389$ . Table 5.9 shows the steps for the path of steepest ascent in the coded units.

When the experiment is completely randomized, runs are typically carried out one at a time on the path of steepest ascent. This is not realistic in the split-plot setting. When setting up the order of the split-plot runs, careful consideration has to be given to the restricted randomization. After a setting in the whole-plot factors is determined, several subplots should be run. One must decide how to run the various whole-plot settings as well as which subplot settings will go in each whole plot. Kowalski et al. (2005) propose three methods for carrying out the method of steepest ascent in a split-plot experiment. We will use the

**TABLE 5.9** Steps in the Coded Variables Along the Path of Steepest Ascent for the Printing Press Example

|                      | $z_1$ | $z_2$ |                      | $x_1$ | $x_2$ | $x_3$ |
|----------------------|-------|-------|----------------------|-------|-------|-------|
| Base1                | 0     | 0     | Base2                | 0     | 0     | 0     |
| $\Delta_i$           | 0.618 | 1.0   | $\Delta_j$           | 0.389 | -1.0  | 0.611 |
| Base1 + $\Delta_i$   | 0.618 | 1.0   | Base2 + $\Delta_j$   | 0.389 | -1.0  | 0.611 |
| Base1 + 2 $\Delta_i$ | 1.236 | 2.0   | Base2 + 2 $\Delta_j$ | 0.778 | -2.0  | 1.222 |
| Base1 + 3 $\Delta_i$ | 1.854 | 3.0   | Base2 + 3 $\Delta_j$ | 1.167 | -3.0  | 1.833 |

**TABLE 5.10 First Whole-Plot Run with Setting Base1+ $\Delta_j$ , Using Method 1**

| Subplot Runs        | $x_1$ | $x_2$ | $x_3$ | Response |
|---------------------|-------|-------|-------|----------|
| Base2 + $\Delta_j$  | 0.389 | -1    | 0.611 | 69       |
| Base2 + $2\Delta_j$ | 0.778 | -2    | 1.222 | 73       |
| Base2 + $3\Delta_j$ | 1.167 | -3    | 1.833 | 75       |
| Base2 + $4\Delta_j$ | 1.556 | -4    | 2.444 | 73       |
| Base2 + $5\Delta_j$ | 1.945 | -5    | 3.055 | 70       |
| Base2 + $6\Delta_j$ | 2.334 | -6    | 3.666 | 66       |

model in Equation 5.16 to illustrate the methods. Also, throughout this discussion, we will use  $m = 1$  for the first step along the whole plot and subplot paths. Typically, there are only 4–6 steps along the path before rate of improvement for the response starts to drop off.

**Method 1** This method fixes a whole plot setting and carries out the subplot path. Then, using the result from the subplot path it carries out the whole plot path with fixed subplot settings. Using  $m = 1$ , the first whole-plot setting is Base1 +  $\Delta_j$ . Within this setting of the whole-plot factors, many subplots can be used for the subplot path. For example, say we use six subplots: Base2 +  $\Delta_j$ , Base2 +  $2\Delta_j$ , ..., Base2 +  $6\Delta_j$  run in random order. This determines the subplot setting with the highest response. Now steps are made along the path in the whole-plot factors. The whole-plot path starts with Base1 +  $\Delta_j$  and Base1 +  $2\Delta_j$ . In each of these whole-plot settings, some replicates, say four, of the highest response setting in the subplot factors from the previous subplot path are run. The average of these four subplot runs serve as the response value for the whole-plot path. Continue running new whole-plot setting along the path until the average response values drop off. This factor setting along the whole-plot path and the previously determined subplot settings from the subplot path could be used as the center for any follow-up experimentation.

Consider the model fit in Equation 5.16. Then, as an example of Method 1, first the runs in Table 5.10 are carried out in random order. The subplot path gives the setting Base2 +  $3\Delta_j$  because it has the highest response (75). Next, more whole-plot settings are used, all with the same subplot setting Base2 +  $3\Delta_j$ . The average response of the subplots will be the whole-plot response. From Table 5.11, the whole-plot response drops off after Base1 +  $2\Delta_j$ . Thus, this example produces the setting of Base1 +  $2\Delta_j$  in the whole-plot variables and the setting of Base2 +  $3\Delta_j$  in the subplot variables as the center of the follow-up experiment. The design point in coded units for this setting is  $z_1 = 1.236$ ,  $z_2 = 2.0$ ,  $x_1 = 1.167$ ,  $x_2 = -3$ , and  $x_3 = 1.833$ .

**TABLE 5.11 New Whole-Plot Runs Using Subplot Setting Base2 +  $3\Delta_j$ , Using Method 1**

| Subplot Setting     | Whole Plot Settings |                    |                    |
|---------------------|---------------------|--------------------|--------------------|
|                     | Base1+ $\Delta_j$   | Base1+ $2\Delta_j$ | Base1+ $3\Delta_j$ |
|                     | Response            | Response           | Response           |
| Base2 + $3\Delta_j$ | 76                  | 79                 | 74                 |
| Base2 + $3\Delta_j$ | 75                  | 77                 | 75                 |
| Base2 + $3\Delta_j$ | 76                  | 77                 | 74                 |
| Base2 + $3\Delta_j$ | 75                  | 76                 | 73                 |
| Average             | 75.5                | 77.25              | 74.0               |

**TABLE 5.12 Whole-Plot Runs with setting Base2 +  $\Delta_j$ , Using Method 2**

| Subplot Setting    | Whole-Plot Settings |                     |                     |
|--------------------|---------------------|---------------------|---------------------|
|                    | Base1 + $\Delta_i$  | Base1 + $2\Delta_i$ | Base1 + $3\Delta_i$ |
| Base2 + $\Delta_j$ | 70                  | 73                  | 71                  |
| Base2 + $\Delta_j$ | 72                  | 75                  | 73                  |
| Base2 + $\Delta_j$ | 69                  | 72                  | 70                  |
| Base2 + $\Delta_j$ | 69                  | 72                  | 70                  |
| Average            | 70.0                | 73.0                | 71.0                |

The number of steps required in this particular example, using the split-plot Method 1 approach, was five. It is important to keep in mind the split-plot nature in determining the steps to termination of the path. In this example, even though 18 runs were carried out, the path terminates after only two steps in the whole-plot variables and three steps in the subplot variables.

**Method 2** Method 2 essentially reverses the roles of whole and subplots from Method 1. Instead of determining the ideal subplot setting first and then finding the whole-plot setting, method two finds the whole-plot setting first and then determines the subplot setting. Therefore, one would start by running two whole-plot settings, Base1 +  $\Delta_i$  and Base1 +  $2\Delta_i$ , with each containing, say, four replicates of Base2 +  $\Delta_j$ . Using the average of the subplot runs as the response, continue running whole-plot settings along the whole-plot path until the average response drops off. This will determine the whole-plot setting. Then, in this whole-plot setting, say, six subplots can be run: Base2 +  $\Delta_j$ , Base2 +  $2\Delta_j$ , ..., Base2 +  $6\Delta_j$ . These can be used to determine where the subplot responses drop off. The setting from this subplot path and the previously determined whole-plot setting from the whole-plot path could serve as the center for any follow-up experimentation.

Using Equation 5.16 to provide an example of Method 2, the runs in Table 5.12 are carried out, resulting in Base1 +  $2\Delta_i$  as the choice for the whole-plot variables. Next, subplot settings are run in random order, all with the same whole-plot setting Base1 +  $2\Delta_i$  from the whole plot path. From Table 5.13, the subplot response drops off after Base2 +  $3\Delta_j$ , so Base1 +  $2\Delta_i$  in the whole-plot variables and Base2 +  $3\Delta_j$  in the subplot variables can be used as the center for any follow-up experimentation. The design point for this setting would be  $z_1 = 1.236$ ,  $z_2 = 2.0$ ,  $x_1 = 1.167$ ,  $x_2 = -3$ , and  $x_3 = 1.833$ . In this example, the same termination point is achieved from Methods 1 and 2, although this will not always be the case.

**TABLE 5.13 Subplot Settings at the Best Whole-Plot Setting, Base1 +  $2\Delta_i$ , Using Method 2**

| Subplot Runs        | $x_1$ | $x_2$ | $x_3$ | Response |
|---------------------|-------|-------|-------|----------|
| Base2 + $\Delta_j$  | 0.389 | -1    | 0.611 | 73       |
| Base2 + $2\Delta_j$ | 0.778 | -2    | 1.222 | 76       |
| Base2 + $3\Delta_j$ | 1.167 | -3    | 1.833 | 78       |
| Base2 + $4\Delta_j$ | 1.556 | -4    | 2.444 | 75       |
| Base2 + $5\Delta_j$ | 1.945 | -5    | 3.055 | 73       |
| Base2 + $6\Delta_j$ | 2.334 | -6    | 3.666 | 69       |

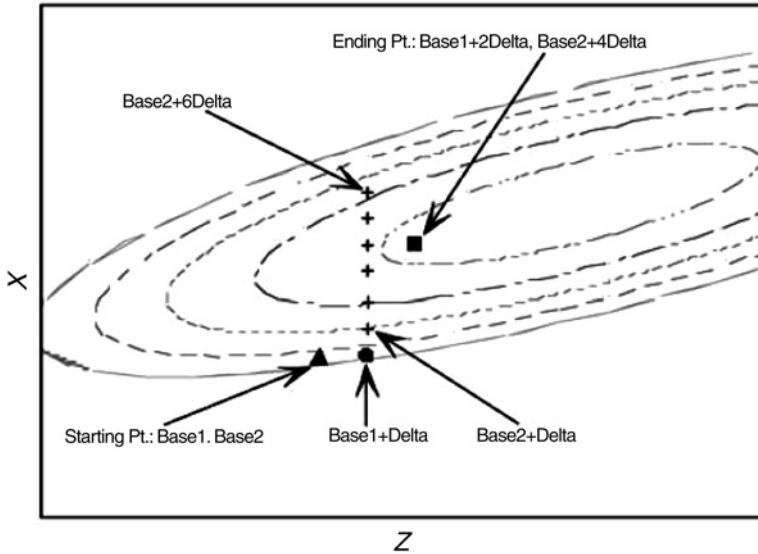


Figure 5.9 Trajectory using Method 1.

It should be noted that both Methods 1 and 2 rely heavily on the assumption of no interaction between whole-plot and subplot factors. However, when the method of steepest ascent is applied in the usual completely randomized RSM experiments, a strict first-order model is assumed to be adequate. Therefore, the assumption of no interaction between whole-plot and subplot factors is not an additional assumption, and is consistent with typical RSM situations.

**Choosing Between Methods 1 and 2** There is no simple way to choose between Methods 1 and 2 in an optimum fashion, where, by optimum, we mean choosing the method that would result in the most rapid movement to the region of the maximum. The optimum choice depends on the starting point in the true response surface and the shape of this surface. These are, of course, unknown to the experimenter. Kowalski et al. (2005) report examining many types of surfaces and used different starting points to study this question empirically. They report that applying Methods 1 and 2 to these surfaces resulted in very little difference in terms of the total number of runs needed. Figures 5.9 and 5.10 illustrate the two methods on one such surface for one starting point (the bold symbols indicate steps along the whole plot path and the + symbol indicates steps along the subplot path). Figure 5.9 uses Method 1, while Fig. 5.10 illustrates Method 2. Both methods require two steps along the whole plot path and 4–5 steps along the subplot path. From a practical point of view, if the largest standardized regression coefficient (coefficient/standard error of the coefficient) belongs to a hard-to-change factor, Method 2 seems like an obvious choice. This would allow the experimenter to make great leaps in the hard-to-change factors toward the optimum and then minor refinements in the easy-to-change factors. On the other hand, if the largest standardized regression coefficient belongs to an easy-to-change factor, Method 1 seems like a better choice because movement in the easy-to-change factors should provide the quicker ascent.

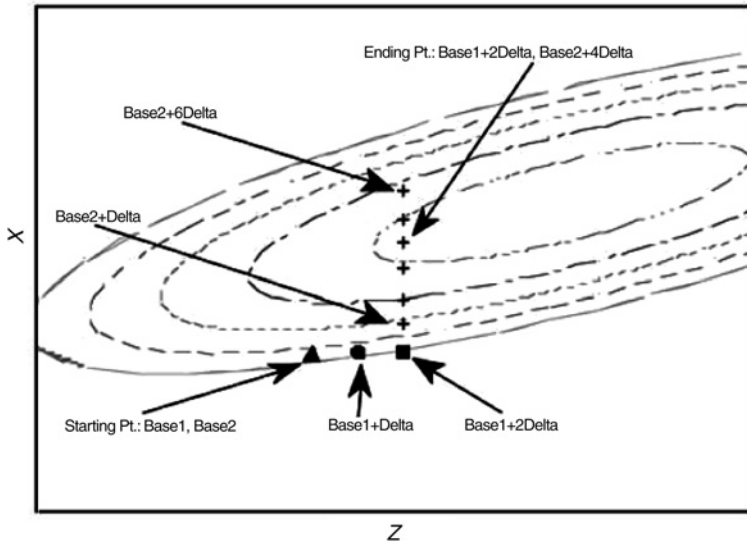


Figure 5.10 Trajectory using Method 2.

It is possible to combine or switch between the two methods. Suppose while using Method 1 there is very little change in the response as the steps along the path in the easy-to-change factors are carried out. This may lead the experimenter to switch to Method 2 and implement steps along the path in the hard-to-change factors to move more quickly to an optimum point.

**Method 3** In this method, we propose running a certain number of whole plots, each with the same number of subplots. This can be thought of as one experiment that carries out both paths at the same time. This approach is preferred if the time to obtain responses is lengthy and collecting data in batches is more time- or cost-effective. Suppose it is practical to run four whole plots, each with six subplots. The whole-plot settings of  $\text{Base1} + \Delta_i$ ,  $\text{Base1} + 2\Delta_i$ ,  $\text{Base1} + 3\Delta_i$ , and  $\text{Base1} + 4\Delta_i$  should be randomized. In each of these settings, the subplot settings of  $\text{Base2} + \Delta_j$ ,  $\text{Base2} + 2\Delta_j, \dots, \text{Base2} + 6\Delta_j$  should be separately randomized for each whole-plot setting. We assume there is no whole-plot by subplot interaction, which is consistent with RSM applications and the model in Equation 5.16. There are two ways to determine the response setting that will serve as the center of a follow-up experiment. One approach is to first find the average of the subplot settings for each whole-plot setting. These can be compared to determine the maximum. Next, the average of the same subplot settings across the whole plots can be calculated to determine the maximum in the subplot settings. Combining these two gives an appropriate point. Another approach is to simply take the highest response from the 24 runs and make this the new center for a follow-up experiment.

Using Equation 5.16 as an example, suppose an experiment with four whole plots and six subplots is chosen, with the resulting runs displayed in Table 5.14. In this example, the highest response occurs at  $(\text{Base1} + 2\Delta_i, \text{Base2} + 3\Delta_j)$ . Also, the highest average subplot setting is  $\text{Base2} + 3\Delta_j$ , and the highest average whole-plot setting is  $\text{Base1} + 2\Delta_i$ . The design point at this setting would be  $z_1 = 1.236$ ,  $z_2 = 2.0$ ,  $x_1 = 1.167$ ,  $x_2 = -3$ , and

**TABLE 5.14 Experimental Runs for Method 3 (Bold Numbers Indicate Highest Responses)**

| Subplots            | Base1 + $\Delta_i$ | Base1 + $2\Delta_i$ | Base1 + $3\Delta_i$ | Base1 + $4\Delta_i$ | Average      |
|---------------------|--------------------|---------------------|---------------------|---------------------|--------------|
| Base2 + $\Delta_j$  | 70                 | 72                  | 72                  | 60                  | 70.25        |
| Base2 + $2\Delta_j$ | 74                 | 75                  | 73                  | 71                  | 73.25        |
| Base2 + $3\Delta_j$ | 74                 | <b>79</b>           | 76                  | 75                  | <b>76.00</b> |
| Base2 + $4\Delta_j$ | 73                 | 75                  | 75                  | 71                  | 73.50        |
| Base2 + $5\Delta_j$ | 70                 | 74                  | 73                  | 70                  | 71.75        |
| Base2 + $6\Delta_j$ | 67                 | 70                  | 70                  | 67                  | 68.50        |
| Average             | 71.33              | <b>74.16</b>        | 73.16               | 70.50               |              |

$x_3 = 1.833$ . This agreement between the highest overall setting and the highest average settings may not occur in all situations. The advantage of Method 3 compared with Methods 1 and 2 is that all the data for the experiment are collected together. This can be important in some industrial settings. For example, in semiconductor manufacturing, multiple runs may be performed on a single-wafer tool during one time period. These runs involve changing both hard-to-change and easy-to-change factors on the tool. The wafers are then sent to a lab for processing, and it may take up to a week to get the response values back. Therefore, it is important to collect all of the data at one time, and it would be impractical to attempt to collect the measurements sequentially, as is required with Methods 1 and 2.

One possible disadvantage of Method 3 is that the response may not drop off after, say, four steps in the hard-to-change factors and/or six steps in the easy-to-change factors. When using any of these methods, it is useful to have a rule for deciding when to stop conducting experiments along the path. The rule of thumb that we have suggested previously is to continue experimenting along the path of steepest ascent until two consecutive runs have resulted in a decrease in the response.

## EXERCISES

**5.1** Given the fitted response function

$$\hat{y} = 72.0 + 3.6x_1 - 2.5x_2$$

which is found to fit well in a localized region of  $(x_1, x_2)$ .

- Plot contours of constant response,  $y$ , in the  $(x_1, x_2)$  plane.
- Calculate and plot the path of steepest ascent generated from this response function. Use  $x_1$  to determine the path, with  $\Delta x_1 = 1$ .

**5.2** Suppose a fitted first-order model is given by

$$y = 25 + 4x_1 + 3x_2 - 2.5x_3$$

Suppose that we wish to maximize the response. Calculate several points along the path of steepest ascent.

**5.3** Reconsider the model in Exercise 5.1. Construct a path of steepest ascent, using  $x_2$  to define the path, with  $\Delta x_2 = 1$ .



**TABLE E5.1** Data for Exercise 5.5

| Treatment Combination | Weight % of Phosphorus Compound |               |       |
|-----------------------|---------------------------------|---------------|-------|
|                       | Replication 1                   | Replication 2 | Total |
| (1)                   | 30.3                            | 28.6          | 58.9  |
| <i>a</i>              | 28.5                            | 31.4          | 59.9  |
| <i>b</i>              | 24.5                            | 25.6          | 50.1  |
| <i>ab</i>             | 25.9                            | 27.2          | 53.1  |
| <i>c</i>              | 24.8                            | 23.4          | 48.2  |
| <i>ac</i>             | 26.9                            | 23.8          | 50.7  |
| <i>bc</i>             | 24.8                            | 27.8          | 52.6  |
| <i>abc</i>            | 22.2                            | 24.9          | 47.1  |
| <i>d</i>              | 31.7                            | 33.5          | 65.2  |
| <i>ad</i>             | 24.6                            | 26.2          | 50.8  |
| <i>bd</i>             | 27.6                            | 30.6          | 58.2  |
| <i>abd</i>            | 26.3                            | 27.8          | 54.1  |
| <i>cd</i>             | 29.9                            | 27.7          | 57.6  |
| <i>acd</i>            | 26.8                            | 24.2          | 51.0  |
| <i>bcd</i>            | 26.4                            | 24.9          | 51.3  |
| <i>abcd</i>           | 26.9                            | 29.3          | 56.2  |
| Total                 | 428.1                           | 436.9         | 865.0 |

- (a) Plot this path as the same graph that you constructed in Example 5.1.
- (b) How different are the two paths that you have constructed?
- 5.4** Consider the injection molding experiment in Example 5.2.
- (a) Use the coded design units to determine how far the point Base + 4 $\Delta$  is from the design center.
- (b) Suppose that we had selected  $x_2$  to express the points on the path of steepest descent with  $\Delta x_2 = 1$ . Recompute the points in Table E5.4. Use the point Base + 4 $\Delta$ , and compute the distance of this point from the design center.
- (c) How different are the two points you have computed?
- 5.5** In a metallurgy experiment it is desired to test the effect of four factors and their interactions on the concentration (percent by weight) of a particular phosphorus compound in costing material. The variables are: *A*, percent phosphorus in the refinement; *B*, percent remelted material; *C*, fluxing time, and *D*, holding time. The four factors are varied in a  $2^4$  factorial experiment with two castings taken at each factor combination. The 32 castings were made in random order, and the data are shown in Table E5.1.
- (a) Build a first-order response function.
- (b) Construct a table of the path of steepest ascent in the coded design variables.
- (c) It is important to constrain the percentage of phosphorus and the percentage of remelted material. In fact, in the metric of the coded variables we obtain

$$x_1 + x_2 = 2.7$$

where  $x_1$  is percent phosphorus and  $x_2$  is percent remelted material. Recalculate the path of steepest ascent subject to the above constraint.

- 5.6 Consider the  $2^2$  factorial design featured in the illustration in Figure 3.1 in Chapter 3. Factor  $A$  is the concentration of a reactant, factor  $B$  is the feed rate, and the response is the viscosity of the output material. As one can tell by the computed effects and the analysis of variance shown, the main effects dominate in the Analysis.
- Fit a first-order model.
  - Compute and plot the path of steepest ascent.
  - Suppose that in the metric of the coded variables, the concentration cannot exceed 3.5 and the feed rate cannot go below  $-2.7$ . Show the **practical** path of steepest ascent—that is, the path that takes account of these constraints.
- 5.7 It is stated in the text that the development of the path of steepest ascent makes use of the assumption that the model is truly first-order in nature. However, even if there is a modest amount of curvature or interaction in the system, the use of steepest ascent can be extremely useful in determining a future experimental region. Suppose that in a system involving  $x_1$  and  $x_2$  the actual model is given by

$$E(y) = 14 + 5x_1 - 10x_2 + 3x_1x_2$$

Assume that  $x_1$  and  $x_2$  are in coded form.

- Show a plot of the path of steepest ascent (based on actual parameters) if the interaction is ignored.
  - Show a plot of the path of steepest ascent for the model with interaction. Note that this path is not linear.
  - Comment on the difference in the two paths.
- 5.8 If the first-order model regression coefficients are scaled to a unit vector ( $\sum_{j=1}^k b_j^2 = 1$ ), then any point on the path of steepest ascent is just a multiple of that unit vector. Use this to answer the following questions regarding the model in Exercise 5.1.
- Is the point  $x_1 = 1.23$  and  $x_2 = -3.75$  on the path of steepest ascent?
  - Is the point  $x_1 = 1.75$  and  $x_2 = -4.0$  on the path of steepest ascent?
  - Is the point  $x_1 = 1.44$  and  $x_2 = -1.0$  on the path of steepest ascent?
  - Is the point  $x_1 = 2.0$  and  $x_2 = -1.5$  on the path of steepest ascent?
- 5.9 Consider Example 3.3 in Chapter 3.
- Using a first-order regression model plot the path of steepest ascent.
  - Indicate expected responses in a table for various points on the path.
  - The test for curvature indicated that no quadratic terms are significant. Suppose that, instead of having center runs, four additional runs were placed on the factorial points (one at each point). Will the steepest ascent be improved? Explain.
  - Refer to part (c) above. Give an argument for improvement that takes the confidence region on the path of steepest ascent into account.

**TABLE E5.2** Data for Exercise 5.12

| Natural Variables |         | Coded Variables |       | Response |
|-------------------|---------|-----------------|-------|----------|
| $\xi_1$           | $\xi_2$ | $x_1$           | $x_2$ |          |
| 80                | 40      | -1              | -1    | 15       |
| 100               | 40      | 1               | -1    | 32       |
| 80                | 60      | -1              | 1     | 25       |
| 100               | 60      | 1               | 1     | 40       |
| 90                | 50      | 0               | 0     | 33       |
| 90                | 50      | 0               | 0     | 27       |
| 90                | 50      | 0               | 0     | 30       |

- 5.10** Consider Example 3.5 in Chapter 3. There are two design variables and four blocks. The experiment involves complete blocks, and it is assumed that there is no interaction between blocks and the design variables.
- Write the linear regression model for this experiment. Assume blocks are a part of the model. Assume no interaction between the design variables.
  - Show that the path of steepest ascent is the same whether or not blocking is involved in the experiment.
  - Suppose the block effects are extremely important. Explain how blocking will improve the precision of the path of steepest ascent.
- 5.11** Consider the fractional factorial experiment illustrated in Exercise 4.9.
- From the analysis, write out a first-order regression model. All factors are quantitative. Assume that the levels are centered and scaled to  $\pm 1$  levels.
  - Construct a table giving the path of steepest ascent for achieving maximum yield.
  - Suppose, for example, that the  $CE$  interaction is extremely important but the analyst does not realize it. How does this effect the computed path in part (b)? Give a qualitative explanation.
- 5.12** Consider the data in Table E5.2. Apply the method of steepest ascent and create an appropriate path.
- 5.13** Suppose the model you fitted in Exercise 5.12 was the following:

$$\hat{y} = 28 + 8x_1 - 4.5x_2$$

(Note that this is not necessarily the model *you* actually found when you worked this problem. It's just one that we will use in *this* question, under the assumption that it is the correct model.)

- Which variable would you use to define the step size along the path of steepest ascent in this model, and why?
- Using the variable you selected in part (a), choose a step size that is large enough to take you to the boundary of the experimental region in that particular direction. Find and show the coordinates of this point on the path of steepest ascent in the coded variables  $x_1$  and  $x_2$ .

**TABLE E5.3 Data for Exercise 5.14**

| Natural Variables |         |         | Coded Variables |       |       | Response |
|-------------------|---------|---------|-----------------|-------|-------|----------|
| $\xi_1$           | $\xi_2$ | $\xi_3$ | $x_1$           | $x_2$ | $x_3$ | $y$      |
| 40                | 10      | 80      | -1              | -1    | -1    | 20       |
| 60                | 10      | 80      | 1               | -1    | -1    | 29       |
| 40                | 30      | 80      | -1              | 1     | -1    | 18       |
| 60                | 30      | 80      | 1               | 1     | -1    | 32       |
| 40                | 10      | 120     | -1              | -1    | 1     | 29       |
| 60                | 10      | 120     | 1               | -1    | 1     | 28       |
| 40                | 30      | 120     | -1              | 1     | 1     | 17       |
| 60                | 30      | 120     | 1               | 1     | 1     | 25       |

- (c) Find and show the coordinates of the point from part (c) in terms of the natural variables.
  - (d) Are the points  $\xi_1 = 107.4$  and  $\xi_2 = 59.8$  on the path of steepest ascent?
- 5.14** Consider the data in Table E5.3, which were observed during an experiment conducted to apply the method of steepest ascent (or descent) to a process.
- (a) Fit a first-order model (without interaction) to these data. Show the model in terms of the coded variables.
  - (b) Which variables would you retain in your model as being important? Justify your answer.
- 5.15** This problem refers to the design used in Exercise 5.14. Assume that the experimenter also included the center runs shown in Table E5.4 in his/her design.
- (a) Find an estimate of  $\sigma$ , the standard deviation of the process.
  - (b) Use the new observations to test for lack of fit of the first-order model. Specifically, does there seem to be an indication of curvature (either interaction or pure second-order effects) present?
- 5.16** Suppose that you have fit the following model

$$\hat{y} = 20 + 10x_1 - 4x_2 + 12x_3$$

- (a) The experimenter decides to choose  $x_1$  as the variable to define the step size for steepest ascent. Do you think that this is a good choice for defining step size? Explain why or why not.

**TABLE E5.4 Center Runs for the Experiment in Table E5.3**

| Natural Variables |         |         | Coded Variables |       |       | Response |
|-------------------|---------|---------|-----------------|-------|-------|----------|
| $\xi_1$           | $\xi_2$ | $\xi_3$ | $x_1$           | $x_2$ | $x_3$ | $y$      |
| 50                | 20      | 100     | 0               | 0     | 0     | 32       |
| 50                | 20      | 100     | 0               | 0     | 0     | 36       |
| 50                | 20      | 100     | 0               | 0     | 0     | 39       |
| 50                | 20      | 100     | 0               | 0     | 0     | 33       |

- (b) The experimenter decides to make the step size exactly one coded unit in the  $x_1$  direction. Find the coordinates of this point from the design center on the path of steepest ascent.
- (c) Suppose that you had used  $x_3$  as the variable to define the direction of steepest ascent. Assuming a step size of one coded unit, find the coordinates of the first step in this direction away from the design center.
- (d) Find the distance between the two points found in parts (b) and (c), respectively.
- 5.17** Suppose that you have fit the following response surface model in terms of coded design variables:

$$\hat{y} = 100 + 10x_1 + 15x_2 - 4x_3$$

The design used to fit this model was a  $2^3$  factorial.

- (a) The experimenter is going to use the method of steepest ascent. He/she decides to make the step size exactly one coded unit in the  $x_2$  direction. Find the coordinates of this point from the design center on the path of steepest ascent in terms of the *coded* variables.
- (b) Suppose that the experimenter had chosen  $x_3$  as the variable to define the step size for steepest ascent. Do you think that this is a good choice for defining step size? Explain why or why not.
- (c) Suppose that the levels of the natural variables  $\xi_1$ ,  $\xi_2$ , and  $\xi_3$  used in the  $2^3$  design were (100, 200), (10, 20), and (50, 80), respectively. What are the coordinates of the point on the path of steepest ascent from part (a) in terms of the *natural* variables?
- (d) Find the coordinates of the unit vector defining the path of steepest ascent.
- (e) Suppose that the experimenters want to run the process at the point (in the natural variables)  $\xi_1 = 204$ ,  $\xi_2 = 23.1$ , and  $\xi_3 = 71.6$ . Assuming that the base point is the design center, is this point on the path of steepest ascent?
- 5.18** Suppose that you have fit the following response surface model in terms of coded design variables:

$$\hat{y} = 250 + 20x_1 + 25x_2 - 8x_3$$

The design used to fit this model was a  $2^3$  factorial.

- (a) The experimenter is using the method of steepest ascent. He/she decides to make the step size exactly one coded unit in the  $x_1$  direction. Find the coordinates of this point from the design center on the path of steepest ascent in terms of *coded* design variables.
- (b) Suppose that the experimenter had chosen  $x_2$  as the variable to define the step size for steepest ascent. Once again, the length of the step is one coded unit. Do you think that this is a better choice for defining step size than was used in part (a)? Explain why or why not.
- (c) Suppose that the levels of the natural variables  $\xi_1$ ,  $\xi_2$ , and  $\xi_3$  used in the  $2^3$  design were (100, 150), (10, 20), and (60, 80), respectively. What are the coordinates

of the point on the path of steepest ascent from part (b) in terms of the *natural* variables?

(d) Find the distance separating the two points that you found in parts (a) and (b).

- 5.19** Suppose that you have fit the following response surface model in terms of coded design variables:

$$\hat{y} = 200 + 25x_1 + 30x_2 - 12x_3$$

The design used to fit this model was a  $2^3$  factorial.

- (a) The experimenter is using the method of steepest ascent. He/she decides to make the step size exactly one coded unit in the  $x_1$  direction. Find the coordinates of this point from the design center on the path of steepest ascent in terms of *coded* design variables.
- (b) Suppose that the experimenter had chosen  $x_2$  as the variable to define the step size for steepest ascent. Once again, the length of the step is one coded unit. Do you think that this is a better choice for defining step size than was used in part (a)? Explain why or why not.
- (c) Suppose that the levels of the natural variables  $\xi_1$ ,  $\xi_2$ , and  $\xi_3$  used in the  $2^3$  design were (100, 200), (10, 12), and (60, 100), respectively. What are the coordinates of the point on the path of steepest ascent from part (b) in terms of the *natural* variables?
- (d) Find the distance separating the two points that you found in parts (a) and (b).
- 5.20** Consider Example 5.2 in the textbook. In this example, the variable with the second-largest coefficient (in absolute value) was used to define the path of steepest ascent. Rework this example using the factor with the largest absolute value regression coefficient ( $x_2$ ).
- (a) Construct a table similar to Table 5.4 in the textbook that shows the first four steps taken along the path, assuming that the step size in coded units is  $\Delta x_2 = 1$ .
- (b) Find the distance separating the point that you have found after the fourth step from part (a) above and the fourth step from Table 5.4 in the text. Use coded units.
- 5.21** Consider the first-order model

$$\hat{y} = 100 + 5x_1 + 8x_2 - 3x_3$$

This model was fit using a  $2^3$  design in the coded variables  $-1 \leq x_i \leq +1$ ,  $i = 1, 2, 3$ . The model fit was adequate. The region of exploration on the natural variables was

$$\begin{aligned} \xi_1 &= \text{temperature (100, 110}^\circ\text{C)} & \xi_2 &= \text{time (1, 2 hr)} \\ \xi_3 &= \text{pressure (50, 75 psi)} \end{aligned}$$

- (a) Which variable would you choose to define the step size along the path of steepest ascent, and why?

- (b) Using the variable in part (a), choose a step size large enough to take you to the boundary of the experimental region in that particular direction. Find and show the coordinates of this point on the path of steepest ascent in the coded variables  $x_i$ .
- (c) Find and show the coordinates of this point on the path of steepest ascent from part (b) using the natural variables.
- (d) Find a unit vector that defines the path of steepest ascent.
- (e) What step size multiplier for the unit vector in (d) above would give the same point on the path of steepest ascent you found in parts (b) and (c)?

**5.22** Consider the fitted first-order model

$$\hat{y} = 15.96 + 1.02x_1 + 3.4x_2 - 2.4x_3$$

where  $x_i$  is a coded variable such that  $-1 \leq x_i \leq 1$ , and the natural variables are

|                            |            |
|----------------------------|------------|
| Pressure( $x_1$ )          | 24–30 psig |
| Time( $x_2$ )              | 14–2 min   |
| Amount of caustic( $x_3$ ) | 2–4 lb     |

If the step size in terms of the natural variable time is 1.5 min, find the coordinates of a point on the path of steepest ascent.

- 5.23** Reconsider the model in Exercise 5.22. Are the following points on the path of steepest ascent?
- (a)  $x_1 = 1.5, x_2 = 3.9, x_3 = -1.8$ .
- (b)  $x_2 = 3.0, x_3 = -1.0$ .
- (c) Pressure = 22 psig, time = 1 min, amount of caustic = 2.2 pounds.
- 5.24** Consider the data of Exercise 5.12. Show graphically a confidence region for the path of steepest ascent.
- 5.25** Reconsider the data of Exercise 5.12. Perform tests for interaction and curvature. From these tests, do you feel comfortable doing the method of steepest ascent? Explain why or why not.
- 5.26** Reconsider the situation described in Exercise 5.1. Suppose that the variance of the regression coefficients is 0.5 and that there are four error degrees of freedom. Find the 95% confidence region on the path of steepest ascent. What fraction of the possible directions from the design origin are excluded by the path of steepest ascent that you have computed?
- 5.27** Rework Exercise 5.26 assuming that the variance of the regression coefficients is 0.25 with four error degrees of freedom. Compare your new results with those obtained previously. What impact does the improved precision of estimation have on the results?
- 5.28** Consider the  $2^5$  design in Exercise 3.38. Fit a model using only the important main effects. Find a path of steepest ascent for this model.

- 5.29 Consider the soup experiment in Exercise 4.31. Fit a main effects model to the data in Table E4.6 and determine an appropriate path of steepest ascent.
- 5.30 Consider the sand casting fractional factorial experiment in Exercise 4.32. The design is shown in Table E4.7. Fit a main effects model in what seems to be the important factors and determine an appropriate path of steepest ascent.
- 5.31 Suppose that you have conducted an experiment using a  $2^3$  factorial experiment and fit the following model in the usual coded units:

$$\hat{y} = 100 + 25x_1 - 14x_2 + 20x_3$$

The two factors  $x_1$  and  $x_2$  are continuous, but  $x_3$  is a two-level categorical factor. How would you apply the method of steepest ascent in this situation?

### Short Answer Questions

- 5.32 Any point on the path of steepest ascent is proportional to the unit vector of the first-order regression coefficients. **True False**
- 5.33 Consider the first-order model  $\hat{y} = 150 + 20x_1 + 12x_2$ . If the experimenter used one coded unit in the  $x_1$  direction as the step size, the point  $x_1 = 1.5, x_2 = 1.0$  is a point on the path of steepest ascent. **True False**
- 5.34 Consider the first-order model in Exercise 5.33. The design used to fit this model was a  $2^2$  factorial with 3 center points. The variance of the predicted response at the point  $x_1 = 1.5, x_2 = 1.0$  is \_\_\_\_\_.
- 5.35 A  $2^3$  design with 4 center points has been used to fit the first-order model  $\hat{y} = 175 + 25x_1 + 10x_2 - 8x_3$ . The variance of the predicted response at any design point other than the center point is \_\_\_\_\_.
- 5.36 When conducting steepest ascent experiments, it may be a good idea to replicate runs at some or all of the points along the path. **True False**
- 5.37 If a first-order model with interaction is used, the path of steepest ascent is a straight line. **True False**
- 5.38 The coordinates of the points on the path of steepest ascent are proportional to the effect estimates from the first-order model. **True False**
- 5.39 For a first-order model with three variables the variance of the predicted response at the design center (in coded units) is  $\sigma^2$ , if an unreplicated  $2^3$  design is used to collect the experimental data. **True False**
- 5.40 The method of steepest ascent is a gradient procedure. **True False**
- 5.41 Suppose that you have run a  $2^3$  design at the start of a steepest ascent experiment and fit the following model:

$$\hat{y} = 15 - 9x_2 + 12x_3$$



The experimenter has discarded  $x_1$  as unimportant and decides to choose  $x_2$  as the variable to define the step size for steepest ascent. Do you think that this is a good choice for defining step size? Briefly explain why or why not.

- 5.42** For the model in Exercise 5.41, the experimenter decides to make the step size exactly one coded unit in the  $x_2$  direction. Find the coordinates of this point from the design center on the path of steepest ascent in terms of the coded variables.
- 5.43** For the model in Exercise 5.41, find the variance of the predicted response at the point on the path of steepest ascent computed in Exercise 5.41.
- 5.44** When using steepest ascent, the usual model that is employed is
- (a) First order
  - (b) First order + interaction
  - (c) Reduced quadratic
  - (d) Full quadratic
- 5.45** A first-order model in four variables has been fit using a  $2^{4-1}$  fractional factorial design ( $n = 8$  runs). The variance of the predicted response at any corner of the design is
- (a)  $5\sigma^2$
  - (b)  $\sigma^2/2$
  - (c)  $5\sigma^2/8$
  - (d)  $\sigma^2/8$
  - (e) None of the above
- 5.46** Orthogonal designs are widely used in steepest ascent experiments. They are variance-optimal designs for
- (a) The first-order model
  - (b) The second-order model
  - (c) Both first- and second-order models
  - (d) None of the above
- 5.47** Adding center points to a two-level design
- (a) Provides an estimate of error
  - (b) Allows estimation of all second-order terms
  - (c) Reduces the variance of all model regression coefficients
  - (d) All of the above
  - (e) None of the above



---

# 6

---

## THE ANALYSIS OF SECOND-ORDER RESPONSE SURFACES

### 6.1 SECOND-ORDER RESPONSE SURFACE

In previous discussions we have noted the importance of matching the choice of designed experiment to our expectation of what we believe the underlying relationship between the factors and response to be. Recall that in many design situations we begin with an assumed model form that is thought to be adequate to describe the shape of the underlying relationship, but may wish to dedicate some runs of the design to evaluating this assumption.

To this point we have confined ourselves to relationships that we believe are well described by models that either are first-order in the design variables or contain first-order plus interaction terms. In Chapters 3 and 4, attention was focused on two-level designs. These designs are natural choices for fitting models containing first-order main effects and low-order interactions. Indeed, the material in Chapter 5 dealing with steepest ascent made use of first-order models only. Nevertheless, we have often mentioned the need for considering and fitting second-order response surface models. For example, in Chapter 3 we discussed the use of multiple center runs as an augmentation of the standard two-level design. Interest centered on the possible existence of curvature beyond what can be captured by the interaction terms. Hence, the inclusion of the center runs allowed for assessment of the adequacy of the model by providing a means to detect the existence of nonzero pure quadratic terms  $b_{11}x_1^2, b_{22}x_2^2, \dots, b_{kk}x_k^2$ . Recall that the use of center runs to augment a standard two-level design allows only a single degree of freedom for the simultaneous estimation of (and thus testing of the existence of) these second-order coefficients. As a result, efficient estimates of  $b_{11}, b_{22}, \dots, b_{kk}$  separately require additional design points.

In Chapter 5, we indicated that in most attempts at product improvement through the gradient technique, called steepest ascent, the investigator will encounter situations where the lack of fit attributable to curvature from these pure second-order terms is found to be

quite significant. In these cases it is likely that the model containing first-order terms and, say, two-factor interaction terms  $b_{12}, b_{13}, \dots, b_{k-1,k}$  is woefully inadequate. As a result, a **second-order response surface model**, discussed in the next section, is a reasonable choice. In addition, there are many situations when the experimenter may believe that the underlying relationship is more complex than a simple first-order or first-order plus interactions model can describe. These are also cases when the use of a second-order model is advisable.

## 6.2 SECOND-ORDER APPROXIMATING FUNCTION

In Chapter 1 the notion of approximating polynomial functions was discussed. In response surface methodology (RSM) work it is assumed that the true functional relationship

$$y = f(\mathbf{x}, \boldsymbol{\theta}) + \varepsilon \quad (6.1)$$

is, in fact, unknown. Here the variables  $x_1, x_2, \dots, x_k$  are in centered and scaled **design units**. The genesis of the first-order approximating model or the model that contains first-order terms and low-order interaction terms is the Taylor series approximation of Equation 6.1. Clearly, a Taylor series expansion of Equation 6.1 through second-order terms would result in a model of the type

$$\begin{aligned} y &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \beta_{11} x_1^2 + \cdots + \beta_{kk} x_k^2 + \beta_{12} x_1 x_2 \\ &\quad + \beta_{13} x_1 x_3 + \cdots + \beta_{k-1,k} x_{k-1} x_k + \varepsilon \\ &= \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j=2}^k \beta_{ij} x_i x_j + \varepsilon \end{aligned} \quad (6.2)$$

The model of Equation 6.2 is the **second-order response surface model**. Note that it contains all model terms through order two. The term  $\varepsilon$  is the usual random error component. The assumptions on the random error are those discussed in Chapters 2, 3, and 4.

### 6.2.1 The Nature of the Second-Order Function and Second-Order Surface

The model of Equation 6.2 is an interesting and very flexible model to describe experimental data in which there is curvature. We do not mean to imply here that all systems containing curvature are well accommodated by this model. There are often times in which a model nonlinear in the parameters is necessary. In some cases one may even require the use of cubic terms (say  $x_1^2 x_2, x_1^3$ , and so on) to achieve an adequate fit. In other cases the curvature may be quite easily handled through the use of transformation on the response itself; for example, in the case where  $k = 2$ ,

$$\ln y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$$

In other situations, transformation of the design variables may lead to a satisfactory approximating function—for example,

$$y = \beta_0 + \beta_1 \ln x_1 + \beta_2 \ln x_2 + \varepsilon$$

At times transformation of both response and design variables may be needed.

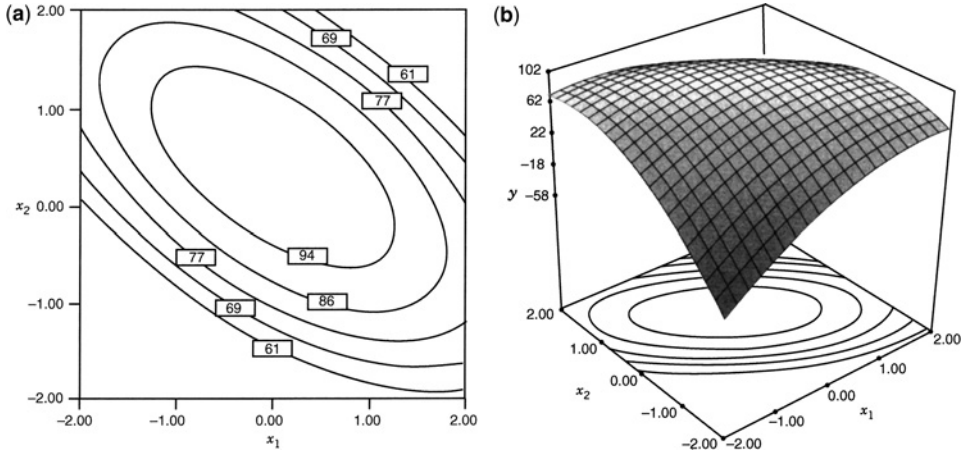


Figure 6.1 Second-order system displaying a maximum. (a) Contour plot. (b) Response surface.

The second-order model described in Equation 6.2 is quite useful and is easily accommodated via the use of a wide variety of experimental designs. Families of designs for fitting the second-order model are discussed in Chapters 8 and 9. Note from Equation 6.2 that the model contains  $1 + 2k + k(k - 1)/2$  parameters. As a result, the experimental design used must contain at least  $1 + 2k + k(k - 1)/2$  distinct design points. In addition, of course, the design must involve at least three levels of each design variable to estimate the pure quadratic terms.

The geometric nature of the second-order function is displayed in Figs. 6.1, 6.2, and 6.3. This will become vital in Section 6.3 as we discuss the location of estimated optimum conditions. Figures 6.1 and 6.2 show contours of constant response for a hypothetical situation with  $k = 2$  variables. In Fig. 6.1 the center of the system, or **stationary point**, is a point of *maximum response*. In Fig. 6.2 the stationary point is a point of *minimum response*. In both cases the response picture displays concentric ellipses. In Fig. 6.3, a hyperbolic

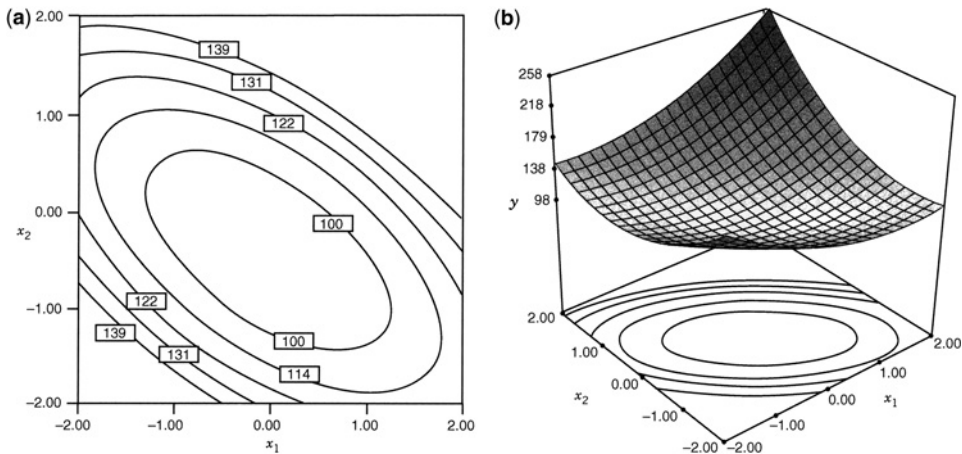
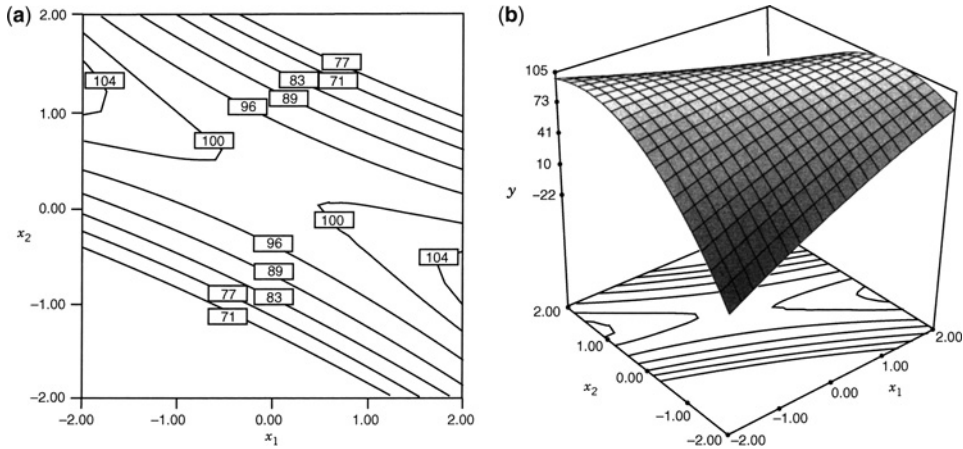


Figure 6.2 Second-order system displaying a minimum. (a) Contour plot. (b) Response surface.



**Figure 6.3** Second-order system displaying a saddle point. (a) Contour plot. (b) Response surface.

system of contours is displayed. Note that the center is neither a maximum nor a minimum point. In this case, the stationary point is called a **saddle point** and the system of contours is called a **saddle** or **minimax system**. The detection of the nature of the system and the location of the stationary point are an important part of the second-order analysis. Modern computer graphics blend nicely with this type of analysis. Three-dimensional graphics can be very helpful to the data analyst in the determination and understanding of the nature of a response surface.

**6.2.2 Illustration of Second-Order Response Surfaces**

The nature of the response surface system (maximum, minimum, or saddle point) depends on the signs and magnitudes of the coefficients in the model of Equation 6.2. The second-order coefficients (interaction and pure quadratic terms) play a vital role. One must keep in mind that the coefficients used are estimates of the  $\beta$ 's of Equation 6.2. As a result, the contours (or surfaces) represent contours of the estimated response. Thus, even the system itself (saddle, maximum, or minimum points) is part of the estimation process and subject to its inherent uncertainty. The stationary point and the general nature of the system arise as a result of a fitted model, not the true structure. For example, if a confidence interval for the estimated quadratic terms of the model contains zero, then the estimated nature of the system could change between a maximum and saddle point system.

Consider the example in Fig. 6.1, for which the fitted second-order model is given by

$$\hat{y} = 100 + 5x_1 + 10x_2 - 8x_1^2 - 12x_2^2 - 12x_1x_2$$

An analysis of this response function would determine the location of the stationary point and the nature of the response system. Because  $k = 2$  in this case, some simple graphics show both issues. Figure 6.1 shows the response surface with contours of constant  $\hat{y}$  plotted against  $x_1$  and  $x_2$ . It shows that the system produces a stationary point that has maximum estimated response. The stationary point is the solution to

$$\frac{\partial \hat{y}}{\partial x_1} = 0 \text{ and } \frac{\partial \hat{y}}{\partial x_2} = 0$$

This results in the system of linear equations

$$\begin{aligned}\frac{\partial \hat{y}}{\partial x_1} &= 0 \rightarrow 16x_1 + 12x_2 = 0 \\ \frac{\partial \hat{y}}{\partial x_2} &= 0 \rightarrow 12x_1 + 24x_2 = 0\end{aligned}$$

yielding the solution for the stationary point

$$x_1 = 0, x_2 = \frac{5}{12}$$

which is the center of the ellipse shown in Fig. 6.1. The estimated response at the stationary point is given by  $\hat{y} = 102.08$ .

In the following sections we discuss a general approach for determining the nature of the response surface contours and the stationary point. Understanding how the estimated equation connects to the nature of the system can be beneficial to interpreting results from the analysis. Graphical summaries, including contour plotting, play an important role in describing the estimated relationship between factors and response. However, there are many times when the formal analysis to describe the nature of the response surface is very helpful. This is particularly true when several ( $k > 2$ ) design variables are involved. Often the nature of the system is elucidated through a combination of analytical techniques and graphical displays. In addition, it is often necessary for the scientist or engineer to use constrained optimization to arrive at potential operating conditions that optimize the response to match study goals. This is particularly true when the stationary point is a saddle point, or when the point of maximum or minimum response resides well outside the experimental region, or when several response variables must be simultaneously considered and optimized. It turns out that the strategy that simultaneously results in process and quality improvement is profoundly dependent on the nature of the response system.

### 6.3 A FORMAL ANALYTICAL APPROACH TO THE SECOND-ORDER MODEL

Consider again the second-order response surface model of Equation 6.2, which we now express with the **fitted model** in matrix notation as

$$\hat{y} = b_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\hat{\mathbf{B}}\mathbf{x} \quad (6.3)$$

where  $b_0$ ,  $\mathbf{b}$ , and  $\hat{\mathbf{B}}$  are the estimates of the intercept, linear, and second-order coefficients, respectively. In fact  $\mathbf{x}' = [x_1, x_2, \dots, x_k]$ ,  $\mathbf{b}' = [b_1, b_2, \dots, b_k]$ , and  $\hat{\mathbf{B}}$  is the  $k \times k$  symmetric matrix

$$\hat{\mathbf{B}} = \begin{bmatrix} b_{11} & b_{12}/2 & \cdots & b_{1k}/2 \\ & b_{22} & \cdots & b_{2k}/2 \\ & & \ddots & \vdots \\ \text{sym.} & & & b_{kk} \end{bmatrix} \quad (6.4)$$

Note that the term  $\mathbf{x}'\hat{\mathbf{B}}\mathbf{x}$  contributes the pure quadratics from the diagonals and the two-way interactions from the sum of two off-diagonal elements  $2(\frac{1}{2}b_{ij}x_i x_j)$ .

### 6.3.1 Location of the Stationary Point

It is straightforward to give a general expression for the location of the stationary point, say  $\mathbf{x}_s$ . One can differentiate  $\hat{y}$  in Equation 6.3 with respect to  $\mathbf{x}$  and obtain

$$\partial\hat{y}/\partial\mathbf{x} = \mathbf{b} + 2\hat{\mathbf{B}}\mathbf{x}$$

Setting the derivative equal to  $\mathbf{0}$ , one can solve for the stationary point of the system:

$$\mathbf{x}_s = -\frac{1}{2}\hat{\mathbf{B}}^{-1}\mathbf{b} \quad (6.5)$$

The **predicted response** at the stationary point is

$$\begin{aligned} \hat{y}_s &= b_0 + \mathbf{x}'_s\mathbf{b} + \mathbf{x}'_s\hat{\mathbf{B}}\mathbf{x}_s \\ &= b_0 + \mathbf{x}'_s\mathbf{b} + \left(-\frac{1}{2}\mathbf{b}'\hat{\mathbf{B}}^{-1}\right)\hat{\mathbf{B}}\mathbf{x}_s \\ &= b_0 + \frac{1}{2}\mathbf{x}'_s\mathbf{b} \end{aligned}$$

### 6.3.2 Nature of the Stationary Point (Canonical Analysis)

The nature of the stationary point is determined from the signs of the eigenvalues of the matrix  $\hat{\mathbf{B}}$ . The relative magnitudes of these eigenvalues can be helpful in the overall interpretation. For example, let the  $k \times k$  matrix  $\mathbf{P}$  be the matrix whose columns are the normalized eigenvectors associated with the eigenvalues of  $\hat{\mathbf{B}}$ . We know that

$$\mathbf{P}'\hat{\mathbf{B}}\mathbf{P} = \Lambda \quad (6.6)$$

where  $\Lambda$  is a diagonal matrix containing the eigenvalues of  $\hat{\mathbf{B}}$  as main diagonal elements. Now if we translate the model of Equation 6.3 to a new center, namely the stationary point, and rotate the axes corresponding to the principal axes of the contour system, we have

$$\begin{aligned} \mathbf{z} &= \mathbf{x} - \mathbf{x}_s \\ \mathbf{w} &= \mathbf{P}'\mathbf{z} \end{aligned} \quad (6.7)$$

The process is illustrated graphically in Fig. 6.4. The translation gives

$$\begin{aligned} \hat{y} &= b_0 + (\mathbf{z} + \mathbf{x}_s)'\mathbf{b} + (\mathbf{z} + \mathbf{x}_s)'\hat{\mathbf{B}}(\mathbf{z} + \mathbf{x}_s) \\ &= [b_0 + \mathbf{x}'_s\mathbf{b} + \mathbf{x}'_s\hat{\mathbf{B}}\mathbf{x}_s] + \mathbf{z}'\mathbf{b} + \mathbf{z}'\hat{\mathbf{B}}\mathbf{z} + 2\mathbf{x}'_s\hat{\mathbf{B}}\mathbf{z} \\ &= \hat{y}_s + \mathbf{z}'\hat{\mathbf{B}}\mathbf{z} \end{aligned}$$

because  $2\mathbf{x}'_s\hat{\mathbf{B}}\mathbf{z} = -\mathbf{z}'\mathbf{b}$  from Equation 6.5. The rotation gives

$$\begin{aligned} \hat{y} &= \hat{y}_s + \mathbf{w}'\mathbf{P}'\hat{\mathbf{B}}\mathbf{P}\mathbf{w} \\ &= \hat{y}_s + \mathbf{w}'\Lambda\mathbf{w} \end{aligned} \quad (6.8)$$



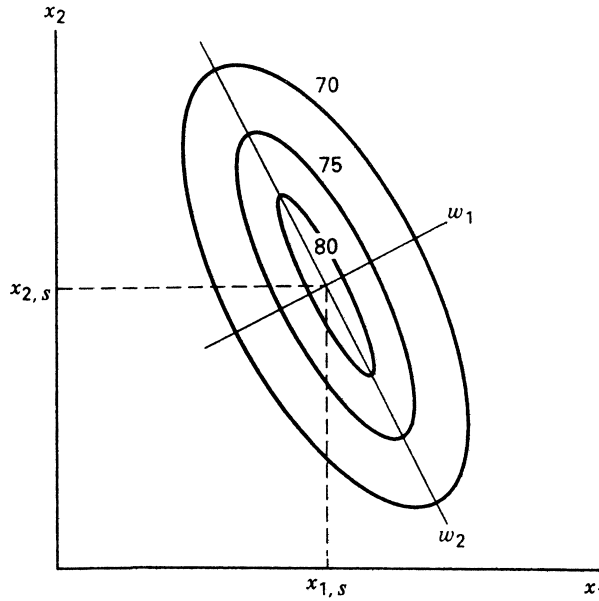


Figure 6.4 Canonical form of the second-order model.

The  $w$ -axes are the principal axes of the contour system. Equation 6.8 can be written

$$\hat{y} = \hat{y}_s + \sum_{i=1}^k \lambda_i w_i^2 \quad (6.9)$$

where  $\hat{y}_s$  is the estimated response at the stationary point and  $\lambda_1, \lambda_2, \dots, \lambda_k$  are the eigenvalues of  $\hat{\mathbf{B}}$ . The variables  $w_1, w_2, \dots, w_k$  are called **canonical variables**.

The translation and rotation described above leads to Equation 6.9. This equation nicely describes the nature of the stationary point and the nature of the system around the stationary point. The *signs* of the  $\lambda$ 's determine the nature of  $\mathbf{x}_s$ , and the relative magnitude of the eigenvalues help the user gain a better understanding of the response system:

1. If  $\lambda_1, \lambda_2, \dots, \lambda_k$  are all *negative*, the stationary point is a point of *maximum* response.
2. If  $\lambda_1, \lambda_2, \dots, \lambda_k$  are all *positive*, the stationary point is a point of *minimum* response.
3. If  $\lambda_1, \lambda_2, \dots, \lambda_k$  are *mixed* in sign, the stationary point is a *saddle* point.

This entire translation and rotation of axes described here is called a **canonical analysis** of the response system. Obviously, if a saddle system occurs, the experimenter will often be led to an alternative type of analysis. However, even if a maximum or minimum resides at  $\mathbf{x}_s$ , the canonical analysis can be helpful. This is true even when one is dealing with a relatively small number of design variables. For example, suppose  $k = 3$  and one seeks a large value for the response. Suppose also that  $\mathbf{x}_s$  is a point of maximum response (all three eigenvalues are negative) but  $|\lambda_1| \cong 0$  and  $|\lambda_2| \cong 0$ . The implication here is that there is considerable flexibility in locating an acceptable set of operating conditions. In fact, there

is essentially an entire plane rather than merely a point where the response is approximately at its maximum value. In fact, it is simple to see that the conditions described by

$$w_3 = 0$$

may in fact describe a rich set of conditions for nearly maximum response. Now, of course,  $w_3 = 0$  describes the plane

$$w_3 = \mathbf{p}'_3 \begin{bmatrix} x_1 - x_{1,s} \\ x_2 - x_{2,s} \\ x_3 - x_{3,s} \end{bmatrix} = 0$$

where the vector  $\mathbf{p}_3$  is the normalized eigenvector associated with the negative eigenvalue  $\lambda_3$ .

**Example 6.1 Canonical Analysis** In Section 2.8 we presented a complete example of fitting a second-order response surface model. The example involved a chemical process in which temperature and concentration were studied, and the response variable was conversion. We used a **central composite design**, originally shown in Table 2.8, but repeated for convenience in Table 6.1. The fitted model is

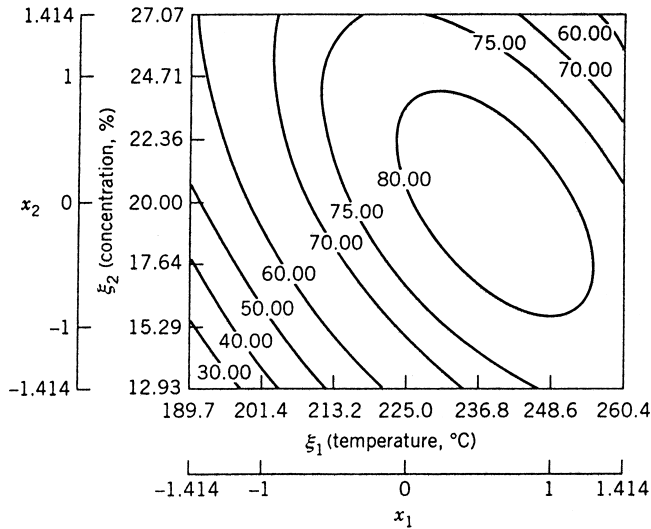
$$\hat{y} = 79.75 + 10.18x_1 + 4.22x_2 - 8.50x_1^2 - 5.25x_2^2 - 7.75x_1x_2$$

and the contour plot for this model is in Fig. 6.5.

*The Central Composite Design* Before we embark on a discussion of the analysis, we should make a few comments about the experimental design. Notice that there are five levels in an experimental design that involves three components. They are: (i) a  $2^2$  factorial, at levels  $\pm 1$  in runs 1–4; (ii) a one-factor-at-a-time array in runs 5–8; and (iii) four

**TABLE 6.1 Central Composite Design for Chemical Process Example**

| Observation | A  |                   | B      |        |     |
|-------------|--|-------------------|--------|--------|-----|
|             | Temperature ( $^{\circ}\text{C}$ ) $\xi_1$ | Conc. (%) $\xi_2$ | $x_1$  | $x_2$  | $y$ |
| 1           | 200  | 15                | -1     | -1     | 43  |
| 2           | 250  | 15                | 1      | -1     | 78  |
| 3           | 200  | 25                | -1     | 1      | 69  |
| 4           | 250  | 25                | 1      | 1      | 73  |
| 5           | 189.65                                     | 20                | -1.414 | 0      | 48  |
| 6           | 260.35                                     | 20                | 1.414  | 0      | 76  |
| 7           | 225  | 12.93             | 0      | -1.414 | 65  |
| 8           | 225  | 27.07             | 0      | 1.414  | 74  |
| 9           | 225  | 20                | 0      | 0      | 76  |
| 10          | 225  | 20                | 0      | 0      | 79  |
| 11          | 225  | 20                | 0      | 0      | 83  |
| 12          | 225  | 20                | 0      | 0      | 81  |



**Figure 6.5** Contour plot of predicted conversion for Example 6.1.

center runs in runs 9–12. A graphical depiction of this design was given in Fig. 2.8. The design essentially consists of eight equally spaced points on a circle of radius  $\sqrt{2}$  and four runs at the design center. The four design points of type (ii) are called **axial points**. This terminology stems from the fact that the points lie on the  $x_1$  or the  $x_2$  axis. Note also that in the axial portion of the design the factors are not varying simultaneously but rather in a one-factor-at-a-time array. As a result, no information regarding the  $x_1x_2$  interaction is provided by this portion of the design. However, the axial portion allows for efficient estimation of pure quadratic terms, that is,  $x_1^2$  and  $x_2^2$ . In Chapter 8, we discuss the central composite design and several other design families in a more thorough way.

*Canonical Analysis* We begin the canonical analysis by determining the location of the stationary point. Note that

$$\hat{\mathbf{B}} = \begin{bmatrix} -8.50 & -3.875 \\ -3.875 & -5.25 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 10.18 \\ 4.22 \end{bmatrix}$$

and from Equation 6.6

$$\begin{aligned} \mathbf{x}_s &= -\frac{1}{2}\mathbf{B}^{-1}\mathbf{b} \\ &= -\frac{1}{2} \begin{bmatrix} -0.1773 & 0.1309 \\ 0.1309 & -0.2871 \end{bmatrix} \begin{bmatrix} 10.18 \\ 4.22 \end{bmatrix} \\ &= \begin{bmatrix} 0.6263 \\ -0.0604 \end{bmatrix} \end{aligned}$$

The predicted response at the stationary point  $(x_1, x_2) = (0.6263, -0.0604)$  is  $\hat{y}_s = 82.81$ . In the natural units, the stationary point is

$$\text{Temperature} = 225 + 25x_{1,s} = 225 + 25(0.6263) = 240.7^\circ\text{C}$$

$$\text{Concentration} = 20 + 5x_{2,s} = 20 + 5(-0.0604) = 19.7\%$$

This is in close agreement with the location of the optimum obtained by visual inspection of Fig. 6.5.

The eigenvalues of the matrix  $\hat{\mathbf{B}}$  are the solutions to the equation

$$|\mathbf{B} - \lambda\mathbf{I}| = 0$$

or

$$\left| \begin{bmatrix} -8.50 - \lambda & -3.875 \\ -3.875 & -5.25 - \lambda \end{bmatrix} \right| = 0$$

This reduces to the quadratic equation

$$\lambda^2 + 13.75\lambda + 29.61 = 0$$

The roots of this equation,  $\lambda_1 = -11.0769$  and  $\lambda_2 = -2.6731$ , are the eigenvalues. Thus the canonical form of the second-order model is

$$\begin{aligned} \hat{y} &= \hat{y}_s + \lambda_1 w_1^2 + \lambda_2 w_2^2 \\ &= 82.81 - 11.0769w_1^2 - 2.6731w_2^2 \end{aligned}$$

Since both eigenvalues are negative, the stationary point is a maximum (as is obvious from inspection of Fig. 6.5). Furthermore, since  $|\lambda_1| > |\lambda_2|$ , the response changes more quickly as we move along the  $w_1$  canonical direction. The surface is somewhat elongated in the  $w_2$  canonical direction with greater distance between the contour lines.

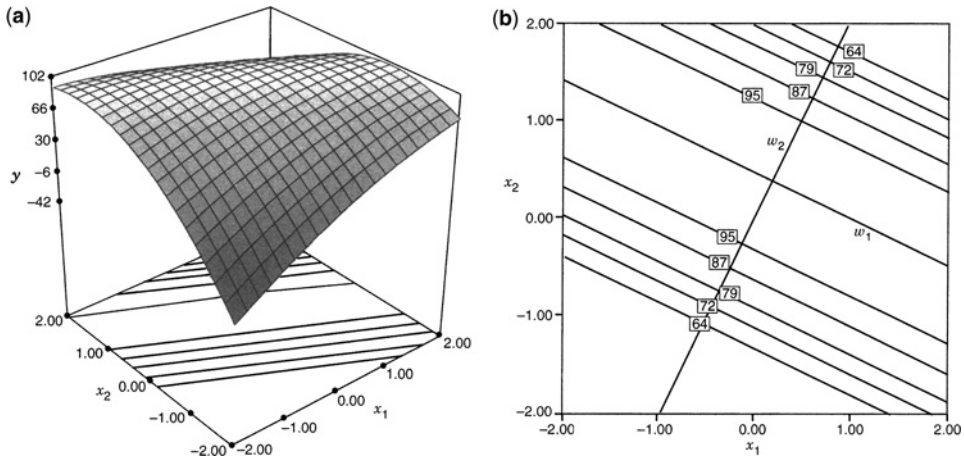
### 6.3.3 Ridge Systems

Variations of the pure minimum, pure maximum, and saddle point surfaces are fairly common. In fact, we suggested such a situation at the conclusion of the previous section, where there were  $k = 3$  variables and two of the eigenvalues were approximately zero ( $|\lambda_1| \cong 0$ ,  $|\lambda_2| \cong 0$ ). A very small (essentially zero) eigenvalue implies considerable elongation of the response surface in that canonical direction, resulting in a **ridge system**.

**Stationary Ridge** Consider the situation shown in Fig. 6.6 for  $k = 2$ , where  $\lambda_1 < 0$  and  $\lambda_2 < 0$  with  $|\lambda_1| \cong 0$ . The canonical model for this surface is

$$\hat{y} = \hat{y}_s + \lambda_2 w_2^2.$$

The stationary point is in the region of the experimental design. This type of response surface is a **stationary ridge system**. The stationary point is a maximum, and yet there is



**Figure 6.6** Stationary ridge system for  $k = 2$ . (a) Response surface. (b) Contour plot.

an approximate maximum on a line *within the design region*. That is, there is essentially a **line maximum**. Consider locations in the design region where  $w_2 = 0$ ; the insensitivity of the response to movement along the  $w_1$ -axis results in little change in response as we move away from the stationary point. The contours of constant response are concentric ellipses that are greatly elongated along the  $w_1$ -axis. Sizable changes in response are experienced along the  $w_2$ -axis.

Stationary ridge systems provide an interesting opportunity when optimizing the response. For this example, if the experimenter wished to maximize the response, there are now many input combinations that provide essentially the same result. Hence, which combination of inputs to select to maximize the response can be chosen based on other considerations. Perhaps some input combinations are less expensive at which to operate, or perhaps, as we will discuss in Chapter 7, there are other responses to simultaneously optimize.

**Rising Ridge** There are other ridge systems that signal certain changes in strategy by the analyst. While the stationary ridge suggests flexibility in choice of operating conditions, the **rising** (or **falling**) ridge suggests that moving outside the experimental region for additional experimentation might warrant consideration. For example, consider a  $k = 2$  situation in which the stationary point is remote from the experimental region. Assume also that one needs to maximize the response. Consider the case where  $\lambda_1$  and  $\lambda_2$  are both negative, but  $\lambda_1$  is near zero. This condition is displayed in Fig. 6.7. The remoteness of the stationary point from the experimental region might suggest here that a movement (experimentally) along the  $w_1$ -axis may well result in an increase in response values. Indeed the rising (or falling) ridge often is a signal to the researcher that he or she has perhaps made a faulty or premature selection of the experimental design region. This is not at all uncommon in applications of RSM. In some cases it may be possible to change the operating conditions to expand to consider factor combinations closer to the stationary point with desirable characteristics.

The rising ridge is signaled by a *stationary point* that is *remote* from the design region. In addition, one of the eigenvalues is near zero. The limiting condition for three variables is

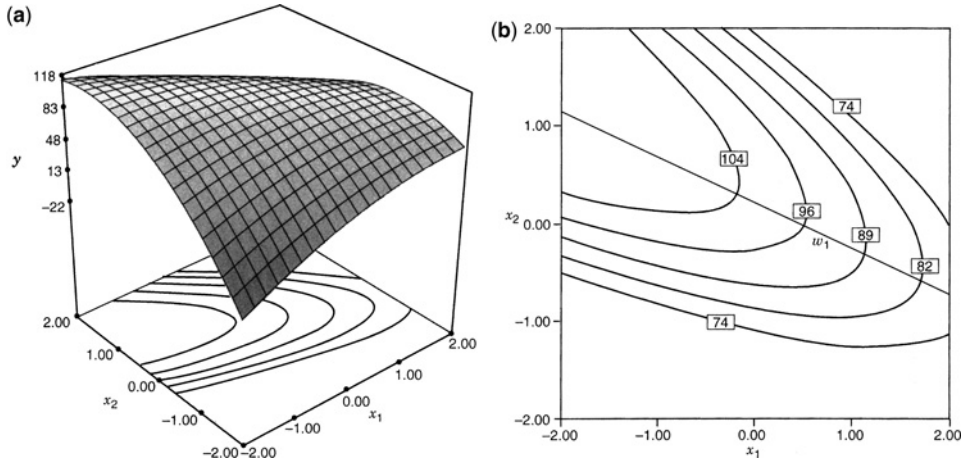


Figure 6.7 Rising ridge system for  $k = 2$ . (a) Response surface. (b) Contour plot.

displayed in the illustration in Fig. 6.8. Here the eigenvalues for  $w_1$ ,  $w_2$ , and  $w_3$  are  $(-, -, \cong 0)$ . Experimentation along the  $w_3$ -axis toward the stationary point is suggested if these represent feasible operating conditions.

Consider the data in Table 6.2 from an experiment on a square region with two factors,  $A$  and  $B$ , using a central composite design ( $2^2$  factorial, four axial runs) with three center runs. The output from Design-Expert is shown in Table 6.3 with the estimated predicted response of

$$\hat{y} = 50.27 - 12.42A + 8.28B - 4.11A^2 - 9.11B^2 + 11.13AB$$

Solving for the canonical form, we find that the response can be re-expressed as

$$\hat{y} = \hat{y}_s - 0.509w_1^2 - 12.71w_2^2$$

where  $x_s = (-5.18, -2.71)$  and the  $w_1$  and  $w_2$  directions are shown in Fig. 6.9b.

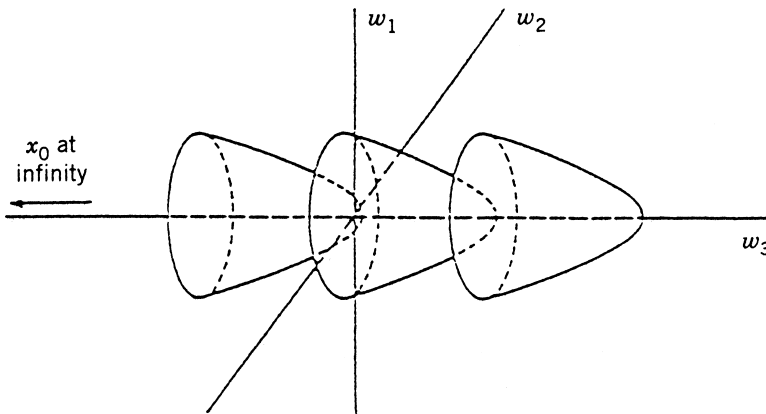


Figure 6.8 Rising ridge conditions in  $k = 3$  variables.

**TABLE 6.2 Design and Data for Rising Ridge System**

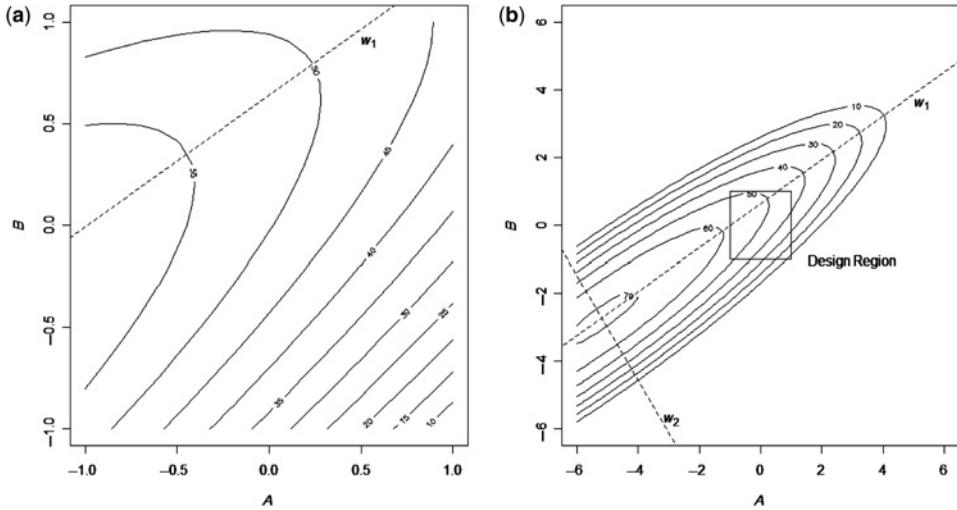
| Standard Order | Observation # | A  | B  | Response |
|----------------|---------------|----|----|----------|
| 1              | 1             | -1 | -1 | 52.3     |
| 2              | 8             | 1  | -1 | 5.3      |
| 3              | 3             | -1 | 1  | 46.7     |
| 4              | 7             | 1  | 1  | 44.2     |
| 5              | 9             | -1 | 0  | 58.5     |
| 6              | 11            | 1  | 0  | 33.5     |
| 7              | 2             | 0  | -1 | 32.8     |
| 8              | 6             | 0  | 1  | 49.2     |
| 9              | 10            | 0  | 0  | 49.3     |
| 10             | 4             | 0  | 0  | 50.2     |
| 11             | 5             | 0  | 0  | 51.6     |

In this case the maximum of the system is located well outside the experimental region, and so it may be helpful to examine both the contour plot of the response for the region where data were observed in Fig. 6.9a and extending the region to include the stationary point in Fig. 6.9b. Note that for each unit moved in the  $w_1$  direction, a 0.509 unit change in the response is observed, while moving a single unit in the  $w_2$  direction results in a 12.71 unit change.

Examples of ridge systems will be seen in illustrations in other portions of this book. In Section 6.3.4 we discuss the role of contour plots in the practical analysis of the second-order fitted surface.

**TABLE 6.3 Design-Expert Output for Rising Ridge System**

| Source  | Sum of Squares       | DF                    | Mean Square    | F-Value                         | Prob > F          |      |
|---|----------------------|-----------------------|----------------|---------------------------------|-------------------|------|
| <i>ANOVA for Response Surface Quadratic Model</i> |                      |                       |                |                                 |                   |      |
| Model   | 2159.04              | 5                     | 431.81         | 750.41                          | <0.0001           |      |
| Residual  | 2.88                 | 5                     | 0.58           |                                 |                   |      |
| <i>Lack of fit</i>                                | 0.20                 | 3                     | 0.068          | 0.051                           | 0.9813            |      |
| <i>Pure Error</i>                                 | 2.67                 | 2                     | 1.34           |                                 |                   |      |
| Cor Total   | 2161.92              | 10                    |                |                                 |                   |      |
| Root MSE  | 0.76                 | <i>R-Squared</i>      |                | 0.9987                          |                   |      |
| Dep Mean  | 43.06                | <i>Adj R-Squared</i>  |                | 0.9873                          |                   |      |
| C.V.  | 1.76                 | <i>Pred R-Squared</i> |                | 0.9970                          |                   |      |
| PRESS   | 6.45                 | <i>Adeq Precision</i> |                | 95.237                          | Desire > 4        |      |
| Factor  | Coefficient Estimate | DF                    | Standard Error | <i>t</i> for $H_0$<br>Coeff = 0 | Prob >   <i>t</i> | VIF  |
| Intercept   | 50.27                | 1                     | 0.39           |                                 |                   |      |
| A-A   | -12.42               | 1                     | 0.31           | -40.09                          | <0.0001           | 1.00 |
| B-B   | 8.28                 | 1                     | 0.31           | 26.75                           | <0.0001           | 1.00 |
| A <sup>2</sup>                                    | -4.11                | 1                     | 0.48           | -8.63                           | 0.0003            | 1.08 |
| B <sup>2</sup>                                    | -9.11                | 1                     | 0.48           | -19.12                          | <0.0001           | 1.08 |
| AB  | 11.13                | 1                     | 0.38           | 29.33                           | <0.0001           | 1.00 |



**Figure 6.9** Contour plot of rising ridge system for  $k = 2$  for data in Table 6.2. (a) Experimental region. (b) Expanded region to include stationary point.

**An Alternate Canonical Form** The canonical form of the second-order model given in Equation 6.8 is usually called the **B-canonical form**. It is very useful for determining the nature of the fitted response surface, particularly in identifying saddle systems and ridges. When the fitted surface is a **rising ridge**, it can be useful to use a slightly different canonical form, namely

$$\hat{y} = a_0 + \mathbf{u}'\mathbf{a} + \mathbf{u}'\Lambda\mathbf{u} \tag{6.10}$$

This is usually called the **A-canonical form**. It basically assumes that the axes have been rotated exactly as in the B-form, but not first translated to the stationary point. In the A-canonical form, the first-order coefficients are

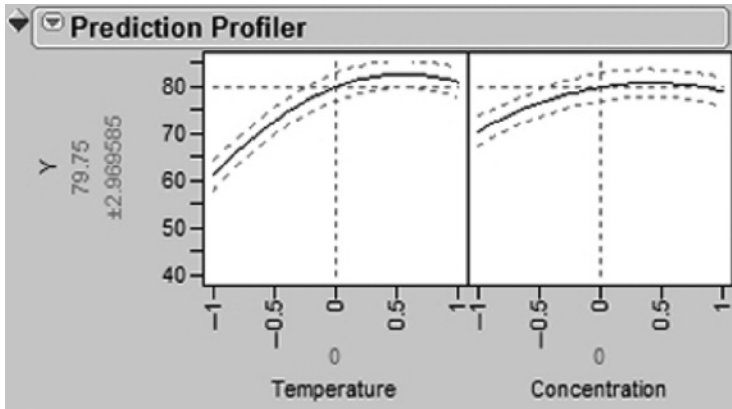
$$\mathbf{a} = \mathbf{P}'\mathbf{b}$$

The first-order terms are useful in determining the direction of movement along the rising ridge towards the optimum.

**6.3.4 Role of Contour Plots**

**Contour plots**, or contour maps, provide one of the most revealing ways of illustrating and interpreting the response surface system. The contour plots are merely two-dimensional (or sometimes three-dimensional) graphs that show contours of constant response with the axis system being a specific pair of the design variables, say  $x_i$  and  $x_j$ , while the other design variables are held constant. Modern graphics allow interesting interpretations that can be of benefit to the user. In fact, in almost all practical response surface situations, the analysis should be followed with a display of contour plotting, as with Example 6.1. The plots are particularly beneficial when the stationary point is not of practical significance (a saddle point, or, say, a maximum point when one seeks a point of minimum response) or when the stationary point is remote from the design region. Clearly, ridge systems can only be





**Figure 6.10** Prediction Profiler in JMP for Example 6.1 response surface.

interpreted when one can observe two-dimensional systems as a set of “snapshots” of the design region. Of course, if there are a large number of design variables, contour plotting is more difficult, since many factors need to be held constant. The method of ridge analysis (Section 6.4) can often aid the user in an understanding of the system.

The prediction profiler in JMP is a dynamic alternative to contour plots. Figure 6.10 illustrates a single view of the slices of the response surface which can be explored using this tool for Example 6.1. Here we have selected scaled temperature = 0 and scaled concentration = 0, but by moving the vertical lines, any combination of temperature and concentration can be examined.

Each plot shows the response surface line with its associated estimation uncertainty conditioned on the other input factor being held constant.

One cautionary note regarding contour plots must be made at this point. The investigator must bear in mind that the contours are only estimates, and if repeated data sets were observed using the same design, the complexion of the response system may change slightly or drastically. In other words, the *contours are not generated by deterministic equations*. Every point on a contour has a standard error. The prediction profiler includes uncertainty bounds on the surface which are helpful reminders that the response surface line is an estimate. This will be illustrated in a more technical way in Section 6.5.

**Example 6.2 Contour Plotting of Survival Data** The following data are adapted from an experiment designed to estimate optimum conditions for storing bovine semen to obtain maximum survival. The variables under study are percent sodium citrate, percent glycerol, and the equilibration time in hours. The response observed is percent survival of motile spermatozoa. The data with the design levels in coded design units are shown in Table 6.4.

The design levels relate to the natural levels in the following way:

|                    | -2  | -1   | 0    | 1    | 2    |
|--------------------|-----|------|------|------|------|
| Sodium citrate     | 1.6 | 2.3  | 3.0  | 3.7  | 4.4  |
| Glycerol           | 2.0 | 5.0  | 8.0  | 11.0 | 14.0 |
| Equilibration time | 4.0 | 10.0 | 16.0 | 22.0 | 28.0 |

TABLE 6.4 Survival Data for Example 6.2

| Treatment Combination | Percent Sodium Citrate | Percent Glycerol | Equilibrium Time (hr) | Percent Survival |
|-----------------------|------------------------|------------------|-----------------------|------------------|
| 1                     | -1                     | -1               | -1                    | 57               |
| 2                     | 1                      | -1               | -1                    | 40               |
| 3                     | -1                     | 1                | -1                    | 19               |
| 4                     | 1                      | 1                | -1                    | 40               |
| 5                     | -1                     | -1               | 1                     | 54               |
| 6                     | 1                      | -1               | 1                     | 41               |
| 7                     | -1                     | 1                | 1                     | 21               |
| 8                     | 1                      | 1                | 1                     | 43               |
| 9                     | 0                      | 0                | 0                     | 60               |
| 10                    | 0                      | 0                | 0                     | 70               |
| 11                    | 0                      | 0                | 0                     | 62               |
| 12                    | 0                      | 0                | 0                     | 72               |
| 13                    | -2                     | 0                | 0                     | 28               |
| 14                    | 2                      | 0                | 0                     | 11               |
| 15                    | 0                      | -2               | 0                     | 2                |
| 16                    | 0                      | 2                | 0                     | 18               |
| 17                    | 0                      | 0                | -2                    | 56               |
| 18                    | 0                      | 0                | 2                     | 46               |

Note that the design consists of a  $2^3$  factorial component, four center runs and six axial or one-factor-at-a-time experimental runs in which each factor is set at level  $+2$  and  $-2$ , while the other factors are set at scaled values of zero. The 18 experimental runs, involving three factors at five evenly spaced levels, are another example of a central composite design. The fitted second-order response function is given by

$$\hat{y} = 66.73 - 1.31x_1 - 2.31x_2 - 1.06x_3 - 11.44x_1^2 - 13.82x_2^2 - 3.57x_3^2 + 9.13x_1x_2 + 0.63x_1x_3 + 0.87x_2x_3$$

From Equation 6.5 the stationary point is given by

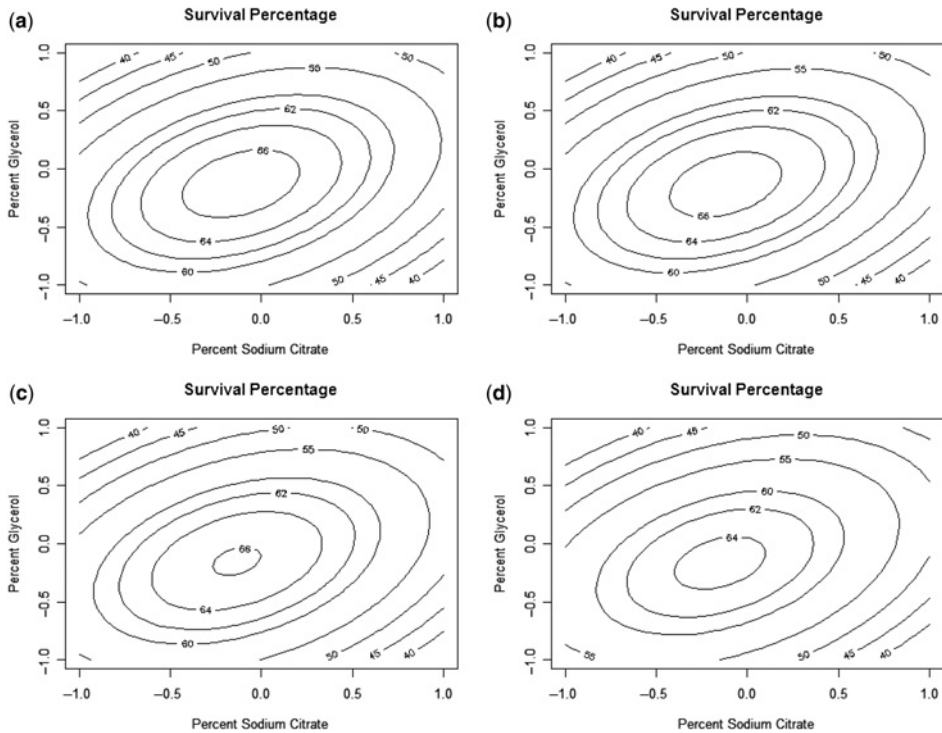
$$\mathbf{x}_s = \begin{bmatrix} x_{1,s} \\ x_{2,s} \\ x_{3,s} \end{bmatrix} = \begin{bmatrix} -0.112 \\ -0.126 \\ -0.174 \end{bmatrix}$$

which is inside the experimental region, and the eigenvalues of the matrix  $\hat{\mathbf{B}}$  are

$$\lambda_1 = -3.508, \quad \lambda_2 = -7.973, \quad \lambda_3 = -17.349$$

The canonical and stationary point analysis indicates that the stationary point is a point of maximum estimated response—that is, *maximum estimated percent survival*. The estimated response at the maximum is  $\hat{y} = 67.041$ . As a result, the estimated condition sought by the experimenters is found at the stationary point inside the experimental design region. Figure 6.11a shows the response surface contours in the vicinity of the optimum with equilibration time 14.96 hr.

Clearly, in this case, the researcher may be quite satisfied to operate at the conditions of the computed stationary point. However, because of economic considerations and scientific

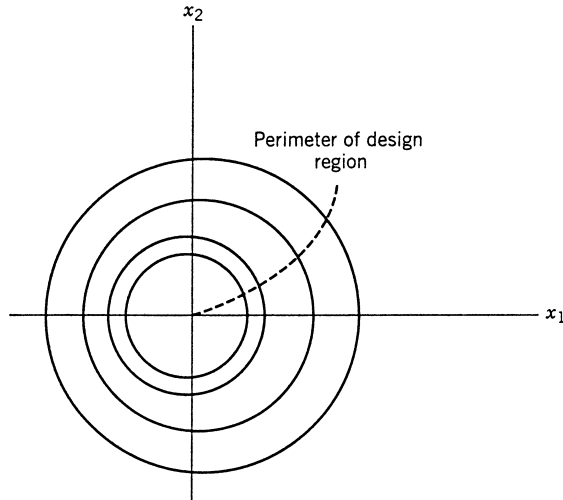


**Figure 6.11** Contours of constant percent survival for Example 6.2. (a) Equilibrium time 14.96 hours. (b) Equilibrium time 14 hours. (c) Equilibrium time 12 hours. (d) Equilibrium time 10 hours.

constraints of the problem, it is often of interest to determine how sensitive the estimated response is to movement away from the stationary point. In this regard, contour plotting can be very useful. For example, here, the stationary point  $\mathbf{x}_s$  corresponds to 2.9% sodium citrate, 7.61 glycerol, and 14.96 hr equilibration time. The equilibration time may be viewed as being somewhat long from a practical point of view, even though the optimum level is inside the experimental design region. Scientific curiosity suggests the need to determine how much is lost in estimated percent survival if equilibration time is set at 14, 12, and 10 hr. An analytical procedure would answer the question, but one gains considerably more information from contour plotting. Figures 6.11b–d reveal the two-dimensional plots giving contours of constant estimated percent survival for fixed values of equilibration times of 14, 12, and 10 hr, respectively. For 14-hr equilibration time, the maximum estimated survival time is over 66%. In fact, Fig. 6.11d reveals that an equilibration time as low as 10 hr can result in an estimated survival time that exceeds 64%. As a result, one could recommend using values of the equilibration time considerably below the optimum and still not sacrifice much in percent survival.

## 6.4 RIDGE ANALYSIS OF THE RESPONSE SURFACE

Often the analyst is confronted with a situation in which the location or nature of the stationary point means that the point is essentially of no value. For example, the stationary point may be a saddle point, which of course is neither a point of estimated maximum



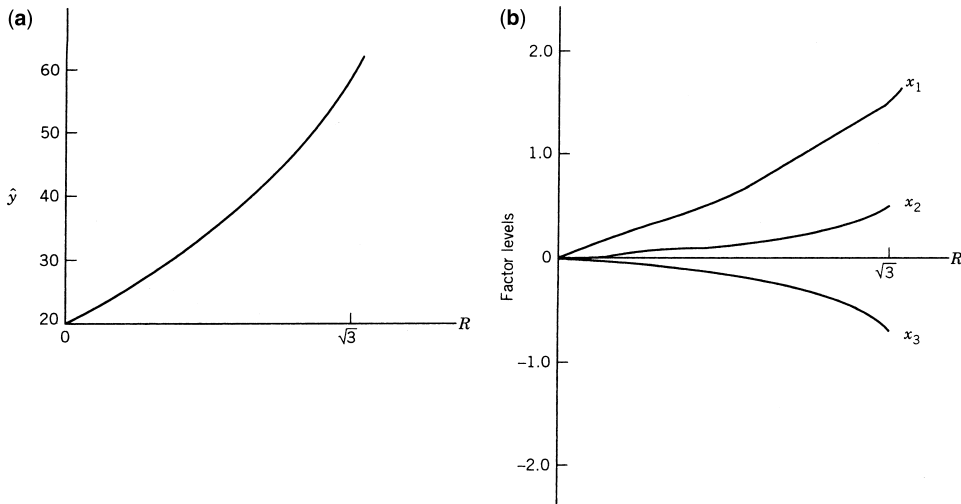
**Figure 6.12** A ridge analysis locus of points to maximize the response for  $k = 2$ .

nor minimum response. Alternately the stationary point might be a point of maximum or minimum response, but well outside the experimental region. In the first case, the stationary point does not represent an important value of the response, while in this second case it would not be advisable to suggest the stationary point as a candidate for operating conditions, because the fitted model is not reliable outside the region of the experiment. The use of contour plotting can be beneficial in these situations, as graphical summaries of the analysis will help describe the nature of the system inside or on the perimeter of the experimental region.

To help formalize information found in contour plots, a useful procedure involves a **constrained optimization** algorithm. It is particularly useful when several design variables are in the response surface model, making visualization of the response values across the entire region difficult. The methodology, called **ridge analysis**, produces a locus of points, each of which is a point of maximum (or minimum, if desired) response, with a constraint that the point resides on the surface of a sphere of a certain radius from the center of the design region. For example, for  $k = 2$  design variables a typical ridge analysis might produce the locus of points as in Fig. 6.12. The origin is  $x_1 = x_2 = 0$ , the design center; a given point, say  $x_p$ , is a point of maximum (or minimum) estimated response, with the constraint being that the point must be on a sphere of radius  $\sqrt{x_p^1 x_p^1 + x_p^2 x_p^2} = r_p$ . In other words, all points identified are constrained optima. Indeed, if the stationary point of the system is not a maximum (or minimum) inside the experimental region, then the point on the path at the design perimeter is the point of maximum (or minimum) estimated response over all points on the path. This, then, may be viewed as a reasonable candidate for recommended operating conditions, if the experimenter is constrained to stay within the design region.

#### 6.4.1 Benefits of Ridge Analysis

The purpose of ridge analysis is to identify a best candidate set of input values constrained to lie inside the experimental region. The output of the analysis is the set of coordinates



**Figure 6.13** Maximum response (a) and constrained optimum conditions (b) plotted against radius value from origin.

of the maxima (or minima) along with the predicted response,  $\hat{y}$ , at each computed point on the path. From this, the analyst gains useful information regarding the roles of the design variables inside the experimental region, as well as potential candidate locations for suggested improved operating conditions. This approach provides a good complement to the contour plots, since with a large number of design variables, many contour plots may be required. Typical output of ridge analysis might simply be a set of two-dimensional plots, regardless of the value of  $k$ .

For example, consider Figs. 6.13a and 6.13b. In this case there are three design variables: Let us assume that the perimeter of the experimental design is at radius  $\sqrt{3}$ . Figure 6.13a shows the maximum value from the fitted response surface for each radius moving away from the origin. Figure 6.13b shows the coordinates of each input for a given radius. If one seeks to maximize the response, a candidate set of conditions is given by  $x_1 \cong 1.5$ ,  $x_2 \cong 0.4$ ,  $x_3 \cong 20.7$ . The estimated maximum response achieved is given by  $\hat{y} \cong 60$ . Another useful feature of ridge analysis is that one obtains some guidelines regarding what input combinations should be considered for future experiments in order to achieve more desirable response values than those obtained in the current experiment. See Marquardt and Snee (1975) for more details.

Ridge analysis as a formal tool is closest to ideal when the design region is spherical. However, even if the design used is not spherical, but rather cuboidal, one can still gain important practical information about the behavior of the response system inside the experimental region, and which design variables are most influential in impacting the response for desired changes in value.

#### 6.4.2 Mathematical Development of Ridge Analysis

In this section we sketch the development of ridge analysis. Many of the details are similar to those of steepest ascent. In fact, *ridge analysis is steepest ascent applied to second-order models*. The reader should recall that the intent of steepest ascent was to provide a path

to an improved product in the case of a system that had not been well studied. A rather inexpensive first-order experiment is the usual vehicle for steepest ascent. However, ridge analysis is generally used when the practitioner feels that he or she is in or quite near the region of the optimum. Nevertheless, both are constrained optimization procedures.

Consider the fitted second-order response surface model of Equation 6.3:

$$\hat{y} = b_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\hat{\mathbf{B}}\mathbf{x}$$

We maximize  $\hat{y}$  subject to the constraint

$$\mathbf{x}'\mathbf{x} = R^2$$

where  $\mathbf{x}' = [x_1, x_2, \dots, x_k]$  and the center of the design region is taken to be  $x_1 = x_2 = \dots = x_k = 0$ . Using Lagrange multipliers, we differentiate

$$L = b_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\hat{\mathbf{B}}\mathbf{x} - \mu(\mathbf{x}'\mathbf{x} - R^2) \tag{6.11}$$

with respect to the vector  $\mathbf{x}$ . The derivative  $\partial L/\partial \mathbf{x}$  is given by

$$\frac{\partial L}{\partial \mathbf{x}} = \mathbf{b} + 2\hat{\mathbf{B}}\mathbf{x} - 2\mu\mathbf{x}$$

To determine constrained stationary points we set  $\partial L/\partial \mathbf{x} = \mathbf{0}$ . This results in

$$(\hat{\mathbf{B}} - \mu\mathbf{I})\mathbf{x} = -\frac{1}{2}\mathbf{b} \tag{6.12}$$

As a result, for a fixed  $\mu$ , a solution  $\mathbf{x}$  of Equation 6.12 is a stationary point on  $R = (\mathbf{x}'\mathbf{x})^{1/2}$ . However, the appropriate solution  $\mathbf{x}$  is that which results in a maximum  $\hat{y}$  on  $R$  or a minimum  $\hat{y}$  on  $R$ , depending on which is desired. It turns out that the appropriate choice of  $\mu$  depends on the eigenvalues of  $\hat{\mathbf{B}}$ . The reader should recall the important role of these eigenvalues in the determination of the nature of the stationary point of the response system (see Section 6.3.2). The following are rules for selection of values of  $\mu$ :

1. If  $\mu$  exceeds the largest eigenvalue of  $\hat{\mathbf{B}}$ , the solution  $\mathbf{x}$  in Equation 6.12 will result in an absolute maximum for  $\hat{y}$  on  $R = (\mathbf{x}'\mathbf{x})^{1/2}$ .
2. If  $\mu$  is smaller the smallest eigenvalue of  $\hat{\mathbf{B}}$ , the solution in Equation 6.12 will result in an absolute minimum for  $\hat{y}$  on  $R = (\mathbf{x}'\mathbf{x})^{1/2}$ .

Consider the orthogonal matrix  $\mathbf{P}$  that diagonalizes  $\mathbf{B}$ . That is,

$$\begin{aligned} \mathbf{P}'\mathbf{B}\mathbf{P} &= \mathbf{\Lambda} \\ &= \begin{bmatrix} \lambda_1 & & 0 & \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_k \end{bmatrix} \end{aligned}$$

where the  $\lambda_i$  are the eigenvalues of  $\mathbf{B}$ . The solution  $\mathbf{x}$  that produces locations where  $\partial L/\partial \mathbf{x} = \mathbf{0}$  is given by

$$(\mathbf{B} - \mu \mathbf{I})\mathbf{x} = -\frac{1}{2}\mathbf{b}$$

If we premultiply  $\mathbf{B} - \mu \mathbf{I}$  by  $\mathbf{P}'$  and postmultiply by  $\mathbf{P}$ , we obtain

$$\mathbf{P}'(\mathbf{B} - \mu \mathbf{I})\mathbf{P} = \mathbf{\Lambda} - \mu \mathbf{I}$$

because  $\mathbf{P}'\mathbf{P} = \mathbf{I}_k$ . If  $\mathbf{B} - \mu \mathbf{I}$  is negative definite, the resulting solution  $\mathbf{x}$  is at least a local maximum on the radius  $R = (\mathbf{x}'\mathbf{x})^{1/2}$ . On the other hand, if  $\mathbf{B} - \mu \mathbf{I}$  is positive definite, the result is a local minimum. Because

$$\begin{aligned} (\mathbf{B} - \mu \mathbf{I}) &= \mathbf{\Lambda} - \mu \mathbf{I} \\ &= \begin{bmatrix} \lambda_1 - \mu & & 0 \\ & \lambda_2 - \mu & \\ 0 & & \ddots \\ & & & \lambda_k - \mu \end{bmatrix} \end{aligned}$$

we see that if  $\mu > \lambda_{\max}$ , then  $\mathbf{B} - \mu \mathbf{I}$  is negative definite, and if  $\mu < \lambda_{\min}$ , then  $\mathbf{B} - \mu \mathbf{I}$  is positive definite.

The above development illustrates but does not prove that the solutions  $\mathbf{x}$  are absolute maxima or minima. For further details see Draper (1963) or Myers (1976).

A computer algorithm for ridge analysis involves the substitution of  $\mu > \lambda_k$  (for a designed maximum response) and increases  $\mu$  until radii near the design perimeter are encountered. Future increases in  $\mu$  results in coordinates that are closer to the design center. The same applies for  $\mu < \lambda_1$  (for desired minimum response), with decreasing values of  $\mu$  being required.

For further discussion and developments related to ridge analysis the reader is referred to Hoerl (1959, 1964), who originated the concept, and Draper (1963), who developed the mathematical ideas behind the concept. The paper by Hoerl (1985) and the book by Box and Draper (2007) provide additional insights into this technique.

**Example 6.3 Ridge Analysis of a Saddle Point Surface** A chemical process that converts 1,2-propanediol to 2,5-dimethylpiperazine is the object of an experiment to determine optimum process conditions—that is, conditions for maximum conversion [see Myers (1976)]. The following factors were studied:

$$\begin{aligned} x_1 &= \frac{\text{amount of NH}_3 - 102}{51} \\ x_2 &= \frac{\text{temperature} - 250}{20} \\ x_3 &= \frac{\text{amount of H}_2\text{O} - 300}{200} \\ x_4 &= \frac{\text{hydrogen pressure} - 850}{350} \end{aligned}$$

As in previous examples, the type of design used for fitting the second-order model was a central composite design. The design matrix and observation vector are as follows:

$$\mathbf{D} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ -1 & -1 & -1 & -1 \\ +1 & -1 & -1 & -1 \\ -1 & +1 & -1 & -1 \\ +1 & +1 & -1 & -1 \\ -1 & -1 & +1 & -1 \\ +1 & -1 & +1 & -1 \\ -1 & +1 & +1 & -1 \\ +1 & +1 & +1 & -1 \\ -1 & -1 & -1 & +1 \\ +1 & -1 & -1 & +1 \\ -1 & +1 & -1 & +1 \\ +1 & +1 & -1 & +1 \\ -1 & -1 & +1 & +1 \\ +1 & -1 & +1 & +1 \\ -1 & +1 & +1 & +1 \\ +1 & +1 & +1 & +1 \\ 0 & 0 & 0 & 0 \\ -1.4 & 0 & 0 & 0 \\ +1.4 & 0 & 0 & 0 \\ 0 & -1.4 & 0 & 0 \\ 0 & +1.4 & 0 & 0 \\ 0 & 0 & -1.4 & 0 \\ 0 & 0 & +1.4 & 0 \\ 0 & 0 & 0 & -1.4 \\ 0 & 0 & 0 & +1.4 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 58.2 \\ 23.4 \\ 21.9 \\ 21.8 \\ 14.3 \\ 6.3 \\ 4.5 \\ 21.8 \\ 46.7 \\ 53.2 \\ 23.7 \\ 40.3 \\ 7.5 \\ 13.3 \\ 49.3 \\ 20.1 \\ 32.8 \\ 31.1 \\ 28.1 \\ 17.5 \\ 49.7 \\ 49.9 \\ 34.2 \\ 31.1 \\ 43.1 \end{bmatrix}$$

The fitted second-order model is given by

$$\begin{aligned}
 \hat{y} = & 40.198 - 1.511x_1 + 1.284x_2 - 8.739x_3 + 4.955x_4 - 6.332x_1^2 \\
 & + 2.194x_1x_2 - 4.292x_2^2 - 0.144x_1x_3 + 8.006x_2x_3 + 0.0196x_3^2 \\
 & + 1.581x_1x_4 + 2.806x_2x_4 + 0.294x_3x_4 - 2.506x_4^2
 \end{aligned}$$

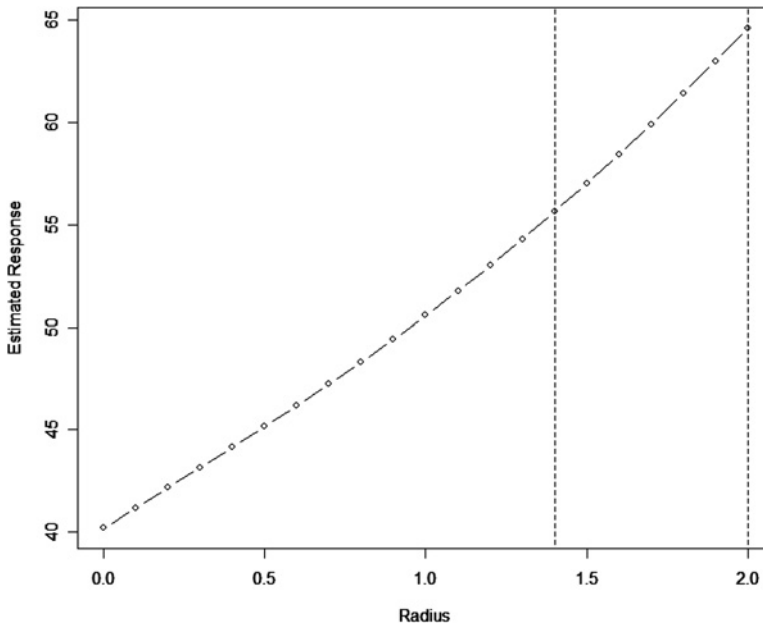
Table 6.5 shows a computer printout of the ridge analysis of maximum response. The accompanying plot of the response versus radius is shown in Fig. 6.14. The stationary point (in design units) is given by  $x_{1,s} = 0.265$ ,  $x_{2,s} = 1.034$ ,  $x_{3,s} = 0.291$ , and  $x_{4,s} = 1.668$ . The response at the stationary point is  $\hat{y}_s = 43.52$ , but the eigenvalues of the matrix  $\hat{\mathbf{B}}$  are  $-7.55$ ,  $-6.01$ ,  $-2.16$ , and  $+2.60$ , indicating a saddle point. They indicate that a ridge analysis might reveal reasonable candidates for operating conditions, with the implied constraint that the candidates lie inside or on the boundary of the experimental region. In this case the factorial points are at a distance of two units from the design center. Note that among the reasonable candidates for coordinates, we have

$$x_1 = -0.1308, \quad x_2 = -0.7861, \quad x_3 = -1.8281, \quad x_4 = 0.1514$$



**TABLE 6.5 Ridge Analysis of Data in Example 6.3**

| Radius | Estimated Response | Standard Error of Prediction | Design Factor Levels |         |         |        |
|--------|--------------------|------------------------------|----------------------|---------|---------|--------|
|        |                    |                              | $x_1$                | $x_2$   | $x_3$   | $x_4$  |
| 0      | 40.198             | 8.321                        | 0                    | 0       | 0       | 0      |
| 0.10   | 41.207             | 8.304                        | -0.0125              | 0.0063  | -0.0870 | 0.0470 |
| 0.20   | 42.195             | 8.254                        | -0.0217              | -0.0012 | -0.1772 | 0.0900 |
| 0.30   | 43.175             | 8.175                        | -0.0287              | -0.0141 | -0.2698 | 0.1270 |
| 0.40   | 44.159             | 8.073                        | -0.0345              | -0.0379 | -0.3640 | 0.1575 |
| 0.50   | 45.157             | 7.960                        | -0.0398              | -0.0685 | -0.4591 | 0.1814 |
| 0.60   | 46.176             | 7.848                        | -0.0450              | -0.1044 | -0.5543 | 0.1993 |
| 0.70   | 47.222             | 7.757                        | -0.0502              | -0.1443 | -0.6493 | 0.2120 |
| 0.80   | 48.301             | 7.707                        | -0.0556              | -0.1873 | -0.7438 | 0.2202 |
| 0.90   | 49.416             | 7.724                        | -0.0611              | -0.2324 | -0.8376 | 0.2247 |
| 1.00   | 50.571             | 7.832                        | -0.0668              | -0.2793 | -0.9307 | 0.2262 |
| 1.10   | 51.767             | 8.055                        | -0.0727              | -0.3274 | -1.0231 | 0.2251 |
| 1.20   | 53.007             | 8.414                        | -0.0787              | -0.3765 | -1.1147 | 0.2219 |
| 1.30   | 54.291             | 8.920                        | -0.0849              | -0.4263 | -1.2057 | 0.2170 |
| 1.40   | 55.621             | 9.581                        | -0.0912              | -0.4767 | -1.2961 | 0.2106 |
| 1.50   | 56.998             | 10.394                       | -0.0976              | -0.5276 | -1.3859 | 0.2029 |
| 1.60   | 58.423             | 11.357                       | -0.1041              | -0.5788 | -1.4752 | 0.1942 |
| 1.70   | 59.896             | 12.461                       | -0.1107              | -0.6303 | -1.5640 | 0.1846 |
| 1.80   | 61.418             | 13.698                       | -0.1173              | -0.6820 | -1.6524 | 0.1742 |
| 1.90   | 62.989             | 15.061                       | -0.1240              | -0.7340 | -1.7404 | 0.1631 |
| 2.00   | 64.610             | 16.543                       | -0.1308              | -0.7860 | -1.8281 | 0.1514 |



**Figure 6.14** Maximum predicted response versus radius for data in Example 6.3.

at a radius of two units. The estimated response is given by  $\hat{y}_s = 64.61$ . Note also that the estimated standard error of prediction at this location is given by

$$s_{\hat{y}} = 16.543$$

This, of course, reflects the variability in  $\hat{y}$  conditioned on the given set of coordinates; that is, given that prediction is desired at the location indicated in  $x_1, x_2, x_3$ , and  $x_4$ . Note how this standard error begins to grow rather rapidly after a radius of 1.4. Recall that the axial distance in the design is 1.4, and the factorial points are at a distance of 2.0. It turns out that a better choice of axial distance would have been 2.0. This would have prevented the quick acceleration of the standard error of prediction as one approaches the design perimeter. In fact, in this case one might recommend conditions at radius 1.4 or 1.5, where the predicted response is smaller but the standard error is considerably smaller. We shed more light on the standard error of prediction in the next section and in future chapters. This criterion plays a critical role in discussions that relate to the choice of experimental design and comparisons among designs.

## 6.5 SAMPLING PROPERTIES OF RESPONSE SURFACE RESULTS

In response surface analysis, considerable focus on optimization and selecting factor combinations has assumed that the results from the analysis are fixed values with no associated uncertainty. In the use of contour plots and the estimation of optimum conditions—either absolute ones, or constrained ones as in the case of ridge analysis—it is important to acknowledge that these are just estimated quantities based on data obtained from an experiment with response values with natural variation. The temptation to offer interpretations that treat these optimum conditions or observed contours of constant response as if the values were scientifically exact can lead to erroneous or misleading conclusions. It is far more helpful to have realistic expectations about what range of response values are possible at the optimum and a region where that optimum is located.

One must bear in mind that in any RSM analysis the computer stationary point is only an estimate, and any point on the contour as well as the contour itself possesses sampling variability. Thus standard errors and confidence regions afford the analyst a more realistic assessment of the quality of the point estimate, and the appropriate amount of uncertainty to associate with it. To gain understanding about how to quantify the uncertainty in predicted values of the response and optimal locations in the design space, we now consider sampling properties of these quantities.

The first type of sampling property we introduce deals with the standard error of predicted response, or standard error of estimated mean response. This concept was discussed briefly in the development of standard regression analysis in Chapter 2.

### 6.5.1 Standard Error of Predicted Response

The most fundamental sampling property in any model building exercise is the standard error of the predicted response at some point of interest  $\mathbf{x}$ , often denoted  $s_{\hat{y}(\mathbf{x})}$ , where in the most general framework

$$\hat{y}(\mathbf{x}) = \mathbf{x}^{(m)'} \mathbf{b}$$

and  $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ . Here, of course, the model matrix  $\mathbf{X}$  is as discussed in Chapter 2. The vector  $\mathbf{x}^{(m)}$  is a function of the location at which one is predicting the response; the  $(m)$  indicates that  $\mathbf{x}^{(m)}$  is just  $\mathbf{x}$  expanded to the **model space**; that is, the vector reflects the form of the model as  $\mathbf{X}$  does. For example, for  $k = 2$  design variables and a second-order model we have

$$\begin{aligned}\mathbf{x}^{(2)'} &= [1, x_1, x_2, x_1^2, x_2^2, x_1x_2] \\ \mathbf{b}' &= [b_0, b_1, b_2, b_{11}, b_{22}, b_{12}]\end{aligned}$$

Under the usual i.i.d. error assumptions for first- and second-order models and with the assumption of constant error variance  $\sigma^2$ , we have

$$\text{Var}[\hat{y}(\mathbf{x})] = \mathbf{x}^{(m)'}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}^{(m)}\sigma^2$$

As a result, an **estimated standard error** of  $\hat{y}(\mathbf{x})$  is given by

$$s_{\hat{y}(\mathbf{x})} = s\sqrt{\mathbf{x}^{(m)'}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}^{(m)}} \quad (6.13)$$

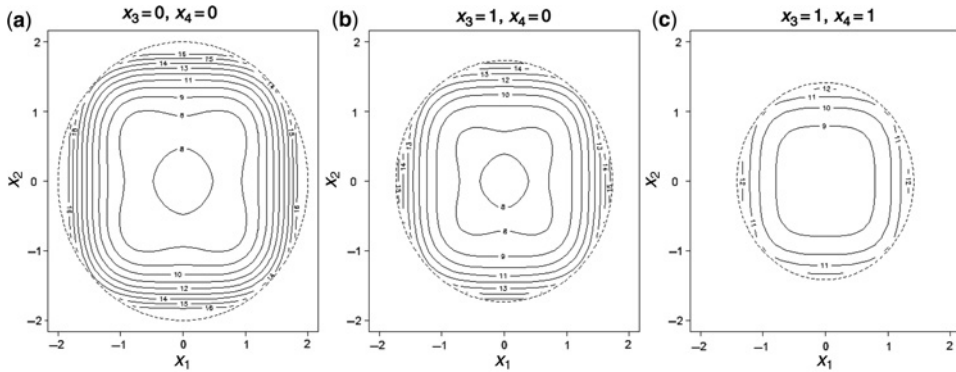
where  $s = \sqrt{MS_E}$  is the square root of the mean square error of the fitted response surface—that is, for a fitted model with  $p$  parameters,

$$s = \sqrt{\sum_{i=1}^k (y_i - \hat{y}_i)^2 / (n - p)}$$

The standard error  $s_{\hat{y}(\mathbf{x})}$  is used in constructing confidence limits around a predicted response. That is, for  $\hat{y}(\mathbf{x}) = \mathbf{x}^{(m)'}\mathbf{b}$ , a prediction at some location  $\mathbf{x}$ , the  $100(1 - \alpha)\%$  confidence interval on the mean response  $E(y|\mathbf{x})$  is given by

$$\hat{y}(\mathbf{x}) \pm t_{\alpha/2, n-p} s \sqrt{\mathbf{x}^{(m)'}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}^{(m)}} \quad (6.14)$$

The standard error  $s_{\hat{y}(\mathbf{x})}$  of the predicted value and the corresponding confidence intervals are the simplest form of sampling properties associated with an RSM analysis. The standard error can provide the user with a rough idea about the relative quality of predicted response values in various locations in the design region. For example, Fig. 6.15 and Table 6.5 show how the standard error of prediction value changes as a function of where we wish to predict in the design space for the ridge system of Example 6.3. Figure 6.15 shows three slices of the design space for fixed values of  $x_3$  and  $x_4$ . Because of the symmetry of the central composite design, the roles of  $x_1, x_2, x_3,$  and  $x_4$  are interchangeable in terms of the standard error of prediction. The dotted lines on each plot indicate a radius of 2 from the center of the design space. Note how prediction with this second-order response surface becomes worse as one gets near the design perimeter. This may very well lead, in some cases, to the conclusion that a reasonable choice of recommended operating conditions might be further inside the perimeter when the predicted value at the optimum on the boundary has a relatively large standard error. The standard error of prediction should be computed at any point that the investigator considers a potentially useful location in the design region.



**Figure 6.15** Contour plots of standard error of predicted value for different  $x_3$  and  $x_4$  values for Example 6.3.

**Standard Error of Prediction as a Function of Design** Because of the prominent role of  $(\mathbf{X}'\mathbf{X})^{-1}$  in Equation 6.13, it should be apparent that  $s_{\hat{y}(\mathbf{x})}$  is very much dependent on the choice of the experimental design. In Table 6.5, the larger standard error that is experienced close to the design perimeter is a result of the choice of design. In the case of the central composite design, the choice of the **axial distance** ( $\pm 1.4$  in this case) and the number of center runs have an important influence on the quality of  $\hat{y}(\mathbf{x})$  as a predictor. In Chapter 8 we deal with properties of first- and second-order designs. Much attention is devoted to  $\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2 = \mathbf{x}^{(m)'}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}^{(m)}$ . In fact, designs are compared on the basis of the relative distribution of values of  $\mathbf{x}^{(m)'}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}^{(m)}$  in the design space.

**What Are the Restrictions on the Standard Error of Prediction?** The user of RSM can learn a great deal about relative quality of prediction using the fitted response surface in various locations in the design space. The prediction standard error—a function of the model, the design, and the location  $\mathbf{x}$ —is quite fundamental. However, it is important for the reader to understand what the standard error of prediction is *not* as well as what it is.

The value  $s_{\hat{y}(\mathbf{x})}$  in Equation 6.13 is an estimate of the standard deviation of  $\hat{y}(\mathbf{x})$  over repeated experiments in which the same design and model are employed, and thus repeated values of  $\hat{y}$  are calculated at the same location  $\mathbf{x}$ .

As a result, if the analyst calculates a stationary point  $\mathbf{x}_s$  (say, a point of maximum estimated response) and determines the standard error of prediction at  $\mathbf{x}_s$ , it cannot be said that the resulting confidence interval is a proper confidence interval on the maximum response. One must keep in mind that repeated experiments would not necessarily produce  $\mathbf{x}_s$ -coordinates at the same location. Indeed, repeated experiments may not all result in stationary points that are maxima. The computed confidence interval here is a confidence interval *conditional* on prediction at the point  $\mathbf{x}_s$ , which only happens to be a stationary point in the present experiment. The confidence region on the location of the stationary point and the confidence interval on the maximum response are considerably more complicated.

In addition, the confidence interval at a given point in the design space is not an interval describing what new observations might be obtained at that combination of inputs. Instead,

the interval describes what the mean of the underlying function is at that location. Recall from Chapter 2 that a  $100(1 - \alpha)\%$  prediction interval for a future observation at the stationary point,  $\mathbf{x}_s$ , can be calculated by adding in the estimated natural variability of the observed responses around the underlying function with an expression of the form,

$$\hat{y}(x_s) \pm t_{\frac{\alpha}{2}, n-p} \sqrt{\hat{\sigma}^2 (1 + \mathbf{x}'_s (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_s)}$$

This prediction interval provides an appropriate range for what values of the response would be obtained at that location, and not necessarily of what values are possible for the overall optimum.

**6.5.2 Confidence Region on the Location of the Stationary Point**

There are instances where the computed stationary point is of considerable interest to the user. It may be the location in the space of the design variables that represents recommended operating conditions for the process. However, there remains the obvious question, “How good are the estimates of these coordinates?” Perhaps there is sufficient uncertainty in the estimator that a secondary, less costly set of conditions results in a mean response that is not significantly different from that produced by the location of the stationary point. These issues can be dealt with nicely through the construction of a confidence region on the location of the stationary point.

Suppose we again consider the fitted second-order response surface model

$$\begin{aligned} \hat{y}(\mathbf{x}) &= b_0 + \sum_{i=1}^k b_i x_i + \sum_{i=1}^k b_{ii} x_i^2 + \sum_{i < j=2}^k b_{ij} x_i x_j \\ &= b_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\hat{\mathbf{B}}\mathbf{x} \end{aligned}$$

One should recall from Section 6.3 that the stationary point is computed from setting the derivatives  $\partial\hat{y}(\mathbf{x})/\partial\mathbf{x}$  to zero. We have

$$\partial\hat{y}(\mathbf{x})/\partial\mathbf{x} = \mathbf{b} + 2\hat{\mathbf{B}}\mathbf{x}$$

The  $j$ th derivative,  $d_j(\mathbf{x})$ , can be written  $d_j(\mathbf{x}) = b_j + 2\hat{\mathbf{B}}'_j\mathbf{x}$ , where the vector  $\hat{\mathbf{B}}'_j$  is the  $j$ th row of  $\hat{\mathbf{B}}$ , the matrix of quadratic coefficients described by Equation 6.5. These derivatives are simple linear functions of  $x_1, x_2, \dots, x_k$ . We denote the vector of derivatives as the  $k$ -dimensional vector  $\mathbf{d}(\mathbf{x})$ . Now, suppose we consider the derivatives evaluated at  $\mathbf{t} = (t_1, t_2, \dots, t_k)$ , where the coordinates of  $\mathbf{t}$  are the coordinates of the *true stationary point of the system* (which of course are unknown). If the errors around the model of Equation 6.2 are i.i.d. normal,  $N(0, \sigma)$ , then

$$\mathbf{d}(\mathbf{t}) \sim N(\mathbf{0}, \text{Var}[\mathbf{d}(\mathbf{t})])$$

where  $\text{Var}[\mathbf{d}(\mathbf{t})]$  is the variance–covariance matrix of  $\mathbf{d}(\mathbf{t})$ . As a result [see Graybill (1976), Myers and Milton (1991)],

$$\frac{\mathbf{d}'(\mathbf{t})[\widehat{\text{Var}}[\mathbf{d}(\mathbf{t})]]^{-1}\mathbf{d}(\mathbf{t})}{k} \sim F_{k, n-p} \tag{6.15}$$

where  $F_{k,n-p}$  is the  $F$ -distribution with  $k$  and  $n - p$  degrees of freedom. It is important to note that  $\text{Var}[\mathbf{d}(\mathbf{t})]$  is a  $k \times k$  matrix that contains the error variance  $\sigma^2$  as a multiplier. In addition,  $\text{Var}[\mathbf{d}(\mathbf{t})]$  is clearly a function of  $\mathbf{t}$ . For example, in the case of two design variables, we have

$$\begin{aligned} d_1(\mathbf{t}) &= b_1 + 2 \left( b_{11}t_1 + \frac{b_{12}}{2}t_2 \right) \\ d_2(\mathbf{t}) &= b_2 + 2 \left( \frac{b_{12}}{2}t_1 + b_{22}t_2 \right) \end{aligned}$$

and

$$\text{Var}[\mathbf{d}(\mathbf{t})] = \begin{bmatrix} \text{Var}[d_1(\mathbf{t})] & \text{Cov}[d_1(\mathbf{t}), d_2(\mathbf{t})] \\ \text{Cov}[d_1(\mathbf{t}), d_2(\mathbf{t})] & \text{Var}[d_2(\mathbf{t})] \end{bmatrix}$$

One can easily observe that elements in this matrix came from  $(\mathbf{X}'\mathbf{X})^{-1}\sigma^2$ , the variance-covariance matrix of regression coefficients. The role of  $t_1$  and  $t_2$  should also be apparent. In Equation 6.15 the “ $\widehat{\phantom{x}}$ ” in  $\{\widehat{\text{Var}}[d(\mathbf{t})]\}$  merely implies that  $\sigma^2$  is replaced by the familiar  $s^2$ , the model error mean square. Now, based on Equation 6.15 we have

$$\Pr \left\{ \mathbf{d}'(\mathbf{t}) \left\{ \widehat{\text{Var}}[\mathbf{d}(\mathbf{t})] \right\}^{-1} \mathbf{d}(\mathbf{t}) \leq kF_{\alpha,k,n-p} \right\} = 1 - \alpha \tag{6.16}$$

where  $F_{\alpha,k,n-p}$  is the upper  $\alpha$ th percentile point of the  $F$ -distribution for  $k$  and  $n - p$  degrees of freedom. Note that while  $\mathbf{t}$  is truly unknown, all other quantities in Equation 6.16 are known. As a result, values  $t_1, t_2, \dots, t_k$  that fall inside the  $100(1-\alpha)\%$  confidence region for the stationary point are those that satisfy the following inequality:

$$\mathbf{d}'(\mathbf{t}) \left\{ \widehat{\text{Var}}[\mathbf{d}(\mathbf{t})] \right\}^{-1} \mathbf{d}(\mathbf{t}) \leq kF_{\alpha,k,n-p} \tag{6.17}$$

This useful and important approach was developed by Box and Hunter (1954).

### 6.5.3 Use and Computation of the Confidence Region on the Location of the Stationary Point

As one can observe from Equation 6.17, the computation of the confidence region is quite simple. However, the display of the confidence region carries with it a clumsiness that is similar to that of the response surface display, particularly when several design variables are involved. For two or three design variables a graphical approach can be very informative. The general size (or volume) of the confidence region can be visualized and the analyst can gain some insight regarding how much flexibility is available in the recommendation of optimum conditions.

**Example 6.4 A Chemical Process with  $k = 2$**  A central composite design was used to develop a response surface relating yield ( $y$ ) to temperature and reaction time. The experimental results are shown in Table 6.6.

**TABLE 6.6 The Central Composite Design for Example 6.4**

| $x_1$ (Temp.)    | $x_2$ (Time) |     | $y$   |     |        |
|------------------|--------------|-----|-------|-----|--------|
| -1               | -1           | -1  | 88.55 |     |        |
| 1                | -1           | -1  | 85.80 |     |        |
| -1               | 1            | 1   | 86.29 |     |        |
| 1                | 1            | 1   | 80.44 |     |        |
| -1.414           | 0            | 0   | 85.50 |     |        |
| 1.414            | 0            | 0   | 85.39 |     |        |
| 0                | -1.414       | 0   | 86.22 |     |        |
| 0                | 1.414        | 0   | 85.70 |     |        |
| 0                | 0            | 0   | 90.21 |     |        |
| 0                | 0            | 0   | 90.85 |     |        |
| 0                | 0            | 0   | 91.31 |     |        |
| Coded levels     | -1.414       | -1  | 0     | 1   | 1.414  |
| Temperature (°C) | 110.86       | 115 | 125   | 135 | 139.14 |
| Time (sec)       | 257.58       | 270 | 300   | 330 | 342.42 |

The fitted second-order model is

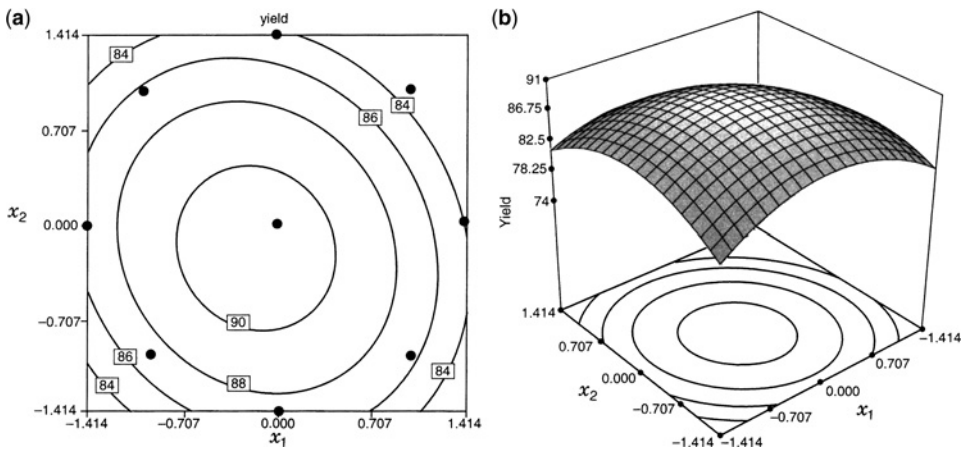
$$\hat{y} = 90.790 - 1.095x_1 - 1.045x_2 - 2.781x_1^2 - 2.524x_2^2 - 0.775x_1x_2$$

with the stationary point at  $\mathbf{x}'_s = [-0.1716, -0.1806]$ . As shown in Fig. 6.16, this is a point of maximum response with the predicted response at the stationary point  $\hat{y}_s = 90.978\%$ . The  $R^2$ -value for this model is approximately 0.84.

The major components in the inequality of Equation 6.17 are quite simple to determine. There are  $p = 6$  parameters in the second-order system, and

$$d_1(\mathbf{t}) = -1.095 - 5.562t_1 - 0.775t_2$$

$$d_2(\mathbf{t}) = -1.045 - 0.775t_1 - 5.048t_2$$



**Figure 6.16** The second-order model for Example 6.4. (a) Contour plot. (b) Response surface.

The variances and covariances are obtained from the matrix  $(\mathbf{X}'\mathbf{X})^{-1}$  with

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} & b_0 & b_1 & b_2 & b_{11} & b_{22} & b_{12} \\ 11 & 0 & 0 & 8 & 8 & 0 \\ & 8 & 0 & 0 & 0 & 0 \\ & & 8 & 0 & 0 & 0 \\ & & & 12 & 4 & 0 \\ & & & & 12 & 0 \\ \text{sym.} & & & & & 4 \end{bmatrix}$$

and

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} \frac{1}{3} & 0 & 0 & -\frac{1}{6} & -\frac{1}{6} & 0 \\ & \frac{1}{8} & 0 & 0 & 0 & 0 \\ & & \frac{1}{8} & 0 & 0 & 0 \\ & & & 0.1772 & 0.0521 & 0 \\ & & & & 0.1772 & 0 \\ \text{sym.} & & & & & \frac{1}{4} \end{bmatrix}$$

The error mean square from the fitted response surface is  $s^2 = 3.1635$  (with five degrees of freedom). Apart from  $s^2$ , the variances and covariances of the coefficients of the fitted response surface are the elements of  $(\mathbf{X}'\mathbf{X})^{-1}$  above.

As a result, elements of  $\widehat{\text{Var}}[\mathbf{d}(t)]$  are easily determined. For example,

$$\widehat{\text{Var}}[d_1(\mathbf{t})] = \frac{s^2}{\sigma^2} \{ \text{Var}b_1 + 4t_1^2 \text{Var}b_{11} + t_2^2 \text{Var}b_{12} \}$$

Note that the covariances involving  $b_1$  are zero and  $\text{Cov}(b_{11}, b_{12}) = \text{Cov}(b_{12}, b_{22}) = 0$ . Thus

$$\begin{aligned} \widehat{\text{Var}}[d_2(\mathbf{t})] &= \frac{s^2}{\sigma^2} \{ \text{Var}b_2 + t_1^2 \text{Var}b_{12} + 4t_2^2 \text{Var}b_{22} \} \\ \widehat{\text{Cov}}[d_1(\mathbf{t}), d_2(\mathbf{t})] &= \frac{s^2}{\sigma^2} \{ 4t_1t_2 \text{Cov}(b_{11}, b_{22}) + t_1t_2 \text{Var}b_{12} \} \end{aligned}$$

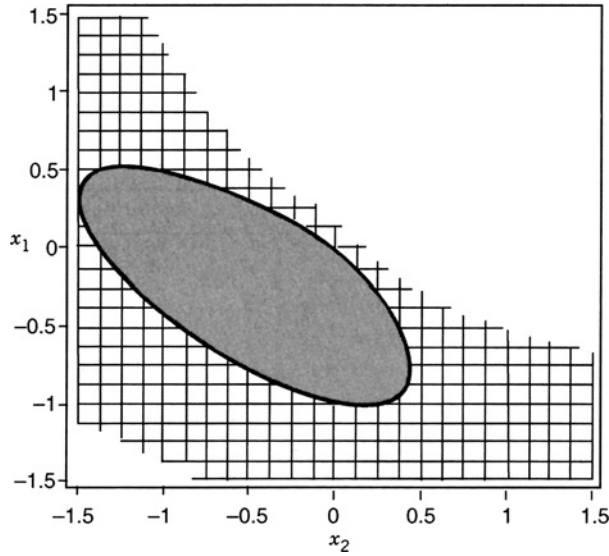
As a result, we have

$$\begin{aligned} \widehat{\text{Var}}[d_1(\mathbf{t})] &= 3.1635 \left\{ \frac{1}{8} + 4t_1^2(0.1772) + \frac{1}{4}t_2^2 \right\} \\ \widehat{\text{Var}}[d_2(\mathbf{t})] &= 3.1635 \left\{ \frac{1}{8} + \frac{1}{4}t_1^2 + 4t_2^2(0.1772) \right\} \\ \widehat{\text{Cov}}[d_1(\mathbf{t}), d_2(\mathbf{t})] &= 3.1635 \left\{ 4(0.0521)t_1t_2 + \frac{1}{4}t_1t_2 \right\} \end{aligned}$$

Equation 6.17 can now be used to plot the confidence region. The reader should note how the design has a profound influence on the important elements of  $(\mathbf{X}'\mathbf{X})^{-1}$ . The confidence region is displayed in Fig. 6.17. This graph was obtained from a Maple program written by a student under the supervision of Dr. Enrique Del Castillo.

The display in Fig. 6.17 presents a rather bleak picture concerning the quality of estimation of the stationary point. The larger, cross-hatched region is the 95% confidence region





**Figure 6.17** Confidence regions on location of stationary point for data of Example 6.4. The solid region is a 90% confidence region while the cross-hatched region corresponds to a 95% confidence region.

on the location of the stationary point, while the smaller, solid region is the 90% confidence region. One must not forget the interpretation, which is that a true response surface maximum at any location inside the 95% contour could readily have produced the data that were observed. The 95% interval does not close inside the experimental design region.

Confidence regions on the location of optima like those shown in Fig. 6.17 are not unusual. The quality of the confidence region depends a great deal on the nature of the design and the fit of the model to the data. Recall that  $R^2$  for the fitted model is approximately 84%. One can note the role of the variances (and covariances of coefficients (and, of course, the error mean square) in the analysis. The appearance of the confidence region shown in Fig. 6.17 should serve as a lesson regarding the emphasis that is put on a *point estimate* of optimum conditions. As we have indicated (and will continue to expound), process improvement should be regarded as a continuing, iterative process, and any optimum found in an experiment should perhaps not be considered as important as what one gains in process improvement and learns about the process in general.

Consider now a second data set with the same design as the one in the previous example, with a considerably better fit. The data are in Table 6.7. The second-order model is given by

$$\hat{y} = 90.7000 + 0.0302x_1 - 0.3112x_2 - 2.0316x_1^2 - 1.8816x_2^2 + 0.5750x_1x_2$$

The value of  $R^2$  is 0.9891.

The stationary point (maximum) is at  $(-0.004373, -0.08339)$  with  $\hat{y}_s = 90.7130$ . One would certainly expect the confidence region to be considerably tighter in this case. Figure 6.18 shows the 95% confidence region around the estimated stationary point. We show the graph using the same scale as in the previous example in order to emphasize the contrast. Clearly, the stationary point in this example is estimated very well. It should be emphasized that a large confidence region is not necessarily bad news. A model that fits

**TABLE 6.7** Data for Second Example Showing Confidence Region on Location of Stationary Point

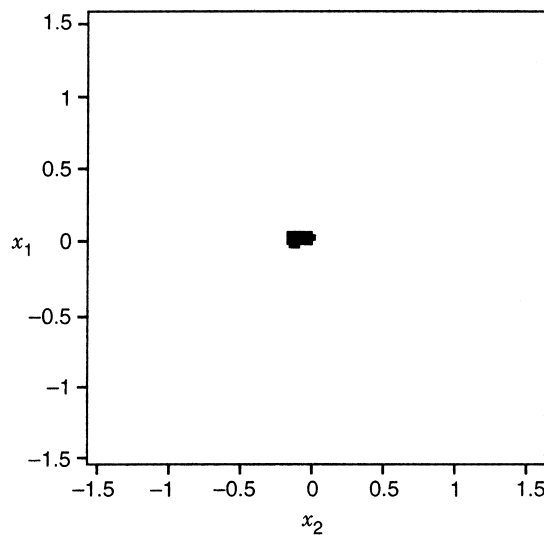
| $x_1$       | $x_2$       | $y$  |
|-------------|-------------|------|
| -1          | -1          | 87.6 |
| -1          | 1           | 85.7 |
| 1           | -1          | 86.5 |
| 1           | 1           | 86.9 |
| $-\sqrt{2}$ | 0           | 86.7 |
| $\sqrt{2}$  | 0           | 86.8 |
| 0           | $-\sqrt{2}$ | 87.4 |
| 0           | $\sqrt{2}$  | 86.7 |
| 0           | 0           | 90.3 |
| 0           | 0           | 91.0 |
| 0           | 0           | 90.8 |

quite well yet generates a large confidence region on the stationary point may produce a “flat” surface, which implies flexibility in choosing the optimum. This is clearly of interest to the engineer or scientist.

The two examples in this section illustrate contrasting situations. One should not get the feeling that the RSM in the first example will not produce important and interesting results about the process. However, in many RSM situations, precise *estimation of optimum conditions is very difficult*.

#### 6.5.4 Confidence Intervals on Eigenvalues in Canonical Analysis

In our discussion of the canonical analysis, we noted that if one or more eigenvalues of the matrix  $\hat{\mathbf{B}}$  are close to zero, a ridge system of some type is present. In judging the size

**Figure 6.18** Confidence region on location of stationary point for data of Table 6.7.

of an eigenvalue it can be helpful to have a *confidence interval* estimate of this parameter. Carter, Chinchilli, and Campbell (1990) and Bisgaard and Ankenman (1996) have described procedures for doing this. Because Bisgaard and Ankenman's procedure is simpler to implement and is equivalent to that of Carter, Chinchilli, and Campbell, we concentrate on it.

Bisgaard and Ankenman's procedure is usually called the **double linear regression (DLR)** method, because it involves fitting an initial second-order model to the data, finding the matrix  $\hat{\mathbf{B}}$  and its eigenvalues and eigenvectors, and then fitting the canonical model in Equation 6.10:

$$\hat{y} = \hat{a}_0 + \sum_{i=1}^k \hat{a}_i u_i + \sum_{i=2}^k \hat{\lambda}_i u_i^2 \quad (6.18)$$

where the variables  $u_1, u_2, \dots, u_k$  are the canonical variables obtained by rotation of the variables  $x_1, x_{21}, \dots, x_k$ :

$$\mathbf{u} = \mathbf{x}'\mathbf{P} \quad (6.19)$$

Because Equation 6.18 is just a standard linear regression model, we can find the standard errors of the  $\hat{\lambda}_i$  directly and obtain a  $100(1 - \alpha)\%$  confidence interval on them using the methods described in Chapter 2, namely,

$$\hat{\lambda}_i \pm t_{\alpha/2, n-p} se(\hat{\lambda}_i)$$

when the model is assumed to have  $p$  parameters.

**Example 6.5 The DLR Method** In Example 6.1 we performed a complete canonical analysis of a second-order model fitted to the chemical process data first introduced in Section 2.8. We found that the eigenvalues of the matrix  $\hat{\mathbf{B}}$  were  $\lambda_1 = -11.0782$  and  $\lambda_2 = -2.6740$ , indicating that the response surface contains a maximum, but it is at least moderately elongated in one direction. We find 95% confidence intervals on the eigenvalues, using the DLR method.

The first step in the DLR procedure is to find the levels of the original coded design variables in the new rotated canonical variables. For Example 6.1 the original design matrix is

$$\mathbf{D} = \begin{bmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ -1.414 & 0 \\ 1.414 & 0 \\ 0 & -1.414 \\ 0 & 1.414 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and from the estimated fitted model

$$\hat{y} = 79.75 + 10.12x_1 + 4.22x_2 - 8.50x_1^2 - 5.25x_2^2 - 7.75x_1x_2$$

we obtain the matrix containing the corresponding eigenvectors

$$\mathbf{P} = \begin{bmatrix} 0.8327 & 0.5538 \\ 0.5538 & -0.8327 \end{bmatrix}$$

Hence, the new canonical-variable design matrix using Equation 6.19 gives

$$\mathbf{U} = \mathbf{DP} = \begin{bmatrix} -1.3864 & -0.2789 \\ 0.2789 & -1.3864 \\ -0.2789 & 1.3864 \\ 1.3864 & 0.2789 \\ -1.1774 & 0.7830 \\ 1.1774 & -0.7830 \\ -0.7830 & -1.1774 \\ 0.7830 & 1.1774 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

The first-order terms of the fitted canonical model are obtained by calculating

$$\mathbf{b}'\mathbf{P} = (10.12, 4.22) \begin{bmatrix} 0.8327 & 0.5538 \\ 0.5538 & -0.8327 \end{bmatrix} = (10.8135, -2.1232)$$

and then fitting Equation 6.18 to obtain

$$\hat{y} = 79.75 + 10.8135u_1 - 2.1232u_2 - 11.0769u_1^2 - 2.6731u_2^2$$

Notice that  $\hat{\lambda}_1$  and  $\hat{\lambda}_2$  are identical to the original eigenvalues of  $\hat{\mathbf{B}}$  (to two decimal places). The standard errors of these estimates are

$$se(\hat{\lambda}_1) = se(\hat{\lambda}_2) = 0.9852$$

and  $t_{0.025,7} = 2.365$ . Therefore the 95% confidence intervals on the eigenvalues are

$$\begin{aligned} -11.0769 - 2.365(0.9852) &\leq \lambda_1 \leq -11.0769 + 2.365(0.9852) \\ -13.4069 &\leq \lambda_1 \leq -8.7469 \end{aligned}$$

and

$$\begin{aligned} -2.6731 - 2.365(0.9852) &\leq \lambda_2 \leq -2.6731 + 2.365(0.9852) \\ -5.003 &\leq \lambda_2 \leq -0.3431 \end{aligned}$$

Because neither confidence interval includes zero, we conclude that both  $\lambda_1$  and  $\lambda_2$  are indeed negative, so the true response surface is a maximum. However, note that the upper confidence limit for  $\lambda_2$  is quite close to zero, implying that there is potentially more elongation in the surface in the  $w_2$ -direction than is visually apparent from Fig. 6.5, and the true response surface could actually be fairly close to a stationary ridge system.

## 6.6 FURTHER COMMENTS CONCERNING RESPONSE SURFACE ANALYSIS

There are several important aspects of RSM that only become apparent after one gains a certain level of experience. In a textbook presentation of RSM, a great deal is taken for granted—the model, the experimental region, the proper choice of response, and, of course, the goal of the study. All, however, need to be handled with care, and each one needs to be a topic of discussion between statistician and scientist or engineer.

The choice of experimental region needs to be made very carefully. Choice of range in the natural variable levels cannot be taken lightly. Ranges that are too small may result in an important factor becoming insignificant in the analysis. This is particularly crucial in the phase of the analysis when variable screening is being done. Often the natural sequential deployment of RSM allows the user to make intelligent and informed choices of variable ranges after preliminary phases of the study have been analyzed. For example, following variable screening, the user should have improved choices of ranges for a steepest ascent procedure. After steepest ascent is accomplished, choice of variable ranges to be used for a second-order analysis should be easier to make.

Let us consider briefly the issue of the goal of the experiment. This textbook develops and emphasizes the need to find optimum conditions. However, one should keep in mind that the estimated optimum conditions from one experiment are indeed only estimates. The next experiment may well find the estimated optimum conditions to be at a different set of coordinates. Also, process technology changes from time to time, and a computed set of optimum conditions may well be fleeting. What is often more important is for the analysis to reveal important information about the process and information about the roles of the variables. The computation of a stationary point, a canonical analysis, or a ridge analysis may well lead to important information and deeper understanding about the process. This in the long run will often be more valuable than a single set of coordinates representing an estimate of optimum conditions.

## EXERCISES

- 6.1** In a study to determine the nature of a response system that relates dry modulus of rupture (psi) in a certain ceramic material with three important independent variables, the following quadratic regression equation was determined [see Hackney and Jones (1969)]:

$$\begin{aligned} \hat{y} \times 10^{-2} = & 6.88 - 0.1466x_1^2 + 0.1875x_1x_2 + 0.2050x_1x_3 + 0.0325x_1 \\ & - 0.0053x_2^2 - 0.1450x_2x_3 + 0.2588x_2 + 0.1359x_3^2 - 0.1363x_3 \end{aligned}$$

The independent variables represent ratios of concentration of various ingredients in the material.

- (a) Determine the stationary point.
- (b) Put the response surface into canonical form, and determine the nature of the stationary point.
- (c) Find the appropriate expressions relating the canonical variables to the independent variables  $x_1$ ,  $x_2$ , and  $x_3$ .
- (d) Generate two-dimensional plots showing contours of constant estimated modulus of rupture. Use  $x_3 = -1, 0, 1$ .

- 6.2** In a chemical engineering experiment dealing with heat transfer in a shallow fluidized bed, data are collected on the following four regressor variables: fluidizing gas flow rate, lb/hr ( $x_1$ ); supernatant gas flow rate, lb/hr ( $x_2$ ); supernatant gas inlet nozzle opening, mm ( $x_3$ ); supernatant gas inlet temperature, °F ( $x_4$ ). The responses measured are heat transfer efficiency ( $y_1$ ) and thermal efficiency ( $y_2$ ). The data are given in Table E6.1.

This is a good example of what often happens in practice. An attempt was made to use a particular second-order design. However, errors in controlling the variables produced a design that is only an approximation of the standard design.

- (a) Center and scale the design variables. That is, create the design matrix in coded form.
- (b) Fit a second-order model for both responses.
- (c) In the case of transfer efficiency ( $y_1$ ), do a canonical analysis and determine the nature of the stationary point. Do the same for thermal efficiency ( $y_2$ ).

**TABLE E6.1 Data for Exercise 6.2**

| Observation | $y_1$   | $y_2$ | $x_1$  | $x_2$  | $x_3$ | $x_4$  |
|-------------|---------|-------|--------|--------|-------|--------|
| 1           | 41.852  | 38.75 | 69.69  | 170.83 | 45    | 219.74 |
| 2           | 155.329 | 51.87 | 113.46 | 230.06 | 25    | 181.22 |
| 3           | 99.628  | 53.79 | 113.54 | 228.19 | 65    | 179.06 |
| 4           | 49.409  | 53.84 | 118.75 | 117.73 | 65    | 281.30 |
| 5           | 72.958  | 49.17 | 119.72 | 117.69 | 25    | 282.20 |
| 6           | 107.702 | 47.61 | 168.38 | 173.46 | 45    | 216.14 |
| 7           | 97.239  | 64.19 | 169.85 | 169.85 | 45    | 223.88 |
| 8           | 105.856 | 52.73 | 169.85 | 170.86 | 45    | 222.80 |
| 9           | 99.438  | 51.00 | 170.89 | 173.92 | 80    | 218.84 |
| 10          | 111.907 | 47.37 | 171.31 | 173.34 | 25    | 218.12 |
| 11          | 100.008 | 43.18 | 171.43 | 171.43 | 45    | 219.20 |
| 12          | 175.380 | 71.23 | 171.59 | 263.49 | 45    | 168.62 |
| 13          | 117.800 | 49.30 | 171.63 | 171.63 | 45    | 217.58 |
| 14          | 217.409 | 50.87 | 171.93 | 170.91 | 10    | 219.92 |
| 15          | 41.725  | 54.44 | 173.92 | 71.73  | 45    | 296.60 |
| 16          | 151.139 | 47.93 | 221.44 | 217.39 | 65    | 189.14 |
| 17          | 220.630 | 42.91 | 222.74 | 221.73 | 25    | 186.08 |
| 18          | 131.666 | 66.60 | 228.90 | 114.40 | 25    | 285.80 |
| 19          | 80.537  | 64.94 | 231.19 | 113.52 | 65    | 286.34 |
| 20          | 152.966 | 43.18 | 236.84 | 167.77 | 45    | 221.72 |

- (d) For the case of transfer efficiency, what levels (natural levels) of the design variables would you recommend if maximum transfer efficiency is sought?
- (e) Do the same for thermal efficiency; that is, find levels of the design variables that you recommend for maximization of thermal efficiency.
- 6.3** A chemical extraction process was studied using a central composite design in an effort to find operating conditions on three variables that maximize yield. The results are shown in Table E6.2.
- (a) Fit a second-order response surface model to these data.
- (b) Use contour plots to find conditions that maximize yield.
- (c) Perform the canonical analysis, including finding the location of the stationary point. Interpret the fitted surface.

**TABLE E6.2** Data for Exercise 6.3

| $x_1$ | $x_2$ | $x_3$ | $y$  |
|-------|-------|-------|------|
| -1    | -1    | -1    | 56.6 |
| 1     | -1    | -1    | 58.5 |
| -1    | 1     | -1    | 48.9 |
| 1     | 1     | -1    | 55.2 |
| -1    | -1    | 1     | 61.8 |
| 1     | -1    | 1     | 63.3 |
| -1    | 1     | 1     | 61.5 |
| 1     | 1     | 1     | 64.0 |
| -1    | 0     | 0     | 61.3 |
| 1     | 0     | 0     | 65.5 |
| 0     | -1    | 0     | 64.6 |
| 0     | 1     | 0     | 65.9 |
| 0     | 0     | -1    | 63.6 |
| 0     | 0     | 1     | 65.0 |
| 0     | 0     | 0     | 62.9 |
| 0     | 0     | 0     | 63.8 |

- 6.4** Rayon whiteness is an important factor for scientists dealing with fabric quality. Whiteness is affected by pulp quality and other processing variables. Some of the variables include: acid bath temperature, °C ( $x_1$ ); cascade acid concentration, % ( $x_2$ ); water temperature, °C ( $x_3$ ); sulfide concentration, % ( $x_4$ ); amount of chlorine bleach, lb/min ( $x_5$ ). Table E6.3 shows an experimental design involving these five design variables. The response is a measure of whiteness.

The coding of the design variables is as follows:

$$x_1 = \frac{\text{temp} - 45}{10}, \quad x_2 = \frac{\text{conc} - 0.5}{0.2}, \quad x_3 = \frac{\text{temp} - 85}{3},$$

$$x_4 = \frac{\text{conc} - 0.25}{0.05}, \quad x_5 = \frac{\text{amt} - 0.4}{0.1}$$

TABLE E6.3 Data for Exercises 6.4

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $y$   |
|-------|-------|-------|-------|-------|-------|
| -1    | -1    | -1    | -1    | -1    | 71.5  |
| 1     | 1     | -1    | -1    | -1    | 76.0  |
| 1     | -1    | 1     | -1    | -1    | 79.9  |
| 1     | -1    | -1    | 1     | -1    | 83.5  |
| 1     | -1    | -1    | -1    | 1     | 89.5  |
| -1    | 1     | 1     | -1    | -1    | 84.2  |
| -1    | 1     | -1    | 1     | -1    | 85.7  |
| -1    | 1     | -1    | -1    | 1     | 94.5  |
| -1    | -1    | 1     | 1     | -1    | 89.4  |
| -1    | -1    | 1     | -1    | 1     | 97.5  |
| -1    | -1    | -1    | 1     | 1     | 103.2 |
| 1     | 1     | 1     | 1     | -1    | 108.7 |
| 1     | 1     | 1     | -1    | 1     | 115.2 |
| 1     | 1     | -1    | 1     | 1     | 111.5 |
| 1     | -1    | 1     | 1     | 1     | 102.3 |
| -1    | 1     | 1     | 1     | 1     | 108.1 |
| -2    | 0     | 0     | 0     | 0     | 80.2  |
| 2     | 0     | 0     | 0     | 0     | 84.1  |
| 0     | -2    | 0     | 0     | 0     | 77.2  |
| 0     | 2     | 0     | 0     | 0     | 85.1  |
| 0     | 0     | -2    | 0     | 0     | 71.5  |
| 0     | 0     | 2     | 0     | 0     | 84.5  |
| 0     | 0     | 0     | -2    | 0     | 77.5  |
| 0     | 0     | 0     | 2     | 0     | 79.2  |
| 0     | 0     | 0     | 0     | -2    | 71.0  |
| 0     | 0     | 0     | 0     | 2     | 90.2  |
| 0     | 0     | 0     | 0     | 0     | 72.1  |
| 0     | 0     | 0     | 0     | 0     | 72.0  |
| 0     | 0     | 0     | 0     | 0     | 72.4  |
| 0     | 0     | 0     | 0     | 0     | 71.7  |
| 0     | 0     | 0     | 0     | 0     | 72.8  |

- (a) Fit an appropriate second-order model in the metric of the coded variables.
- (b) Find the stationary point, and describe the nature of the surface.
- (c) Give a recommendation for operating conditions on the design variables. It is important to maximize whiteness.
- (d) Compute the standard error of prediction at the following design locations:
- (i) Design center (0,0,0,0,0)
  - (ii) All factorial points
  - (iii) All axial points
- (e) Comment on the relative stability of the standard error of prediction.

**6.5** A client from the Department of Mechanical Engineering at Virginia Polytechnic Institute and State University asked for help in analyzing an experiment dealing with gas turbine generators. The voltage output of a generator was measured at various



**TABLE E6.4** Data for Exercise 6.5

| $y$ (volts) | Speed $x_1$ (in./sec) | Extension $x_2$ (in.) |
|-------------|-----------------------|-----------------------|
| 1.23        | 5300                  | 0.000                 |
| 3.13        | 8300                  | 0.000                 |
| 1.22        | 5300                  | 0.012                 |
| 1.92        | 8300                  | 0.012                 |
| 2.02        | 6800                  | 0.000                 |
| 1.51        | 6800                  | 0.012                 |
| 1.32        | 5300                  | 0.006                 |
| 2.62        | 8300                  | 0.006                 |
| 1.65        | 6800                  | 0.006                 |
| 1.62        | 6800                  | 0.006                 |
| 1.59        | 6800                  | 0.006                 |

combinations of blade speed and voltage measuring sensor extension. The data are given in Table E6.4.

- (a) Write the design matrix in coded form.
  - (b) What type of design has been used?
  - (c) Fit an appropriate model.
  - (d) Find the stationary point, and interpret the fitted surface.
  - (e) It is important to determine what value of speed and extension results in a response that exceeds 2.8 volts. Describe where in the design region this voltage can be achieved. Use a two-dimensional contour plot to answer the question.
- 6.6** In Exercise 3.9 we discussed an application dealing with effects on cracking of a titanium alloy. An additional study was made in which the same factors were studied except the heat treatment method. The “high” or “+” method was used and held constant. However, in this case, a second-order design was used so the analyst could fit a second-order model involving curvature. The three factors are pouring temperature  $x_1$ ; titanium content  $x_2$ ; and amount of grain refiner,  $x_3$ . Table E6.5 gives the design points and the response data. The response is the crack length (in millimeters) induced in a sample of the alloy.
- (a) Fit an appropriate model with the data of the above central composite design.
  - (b) From the fitted model in (a), estimate the conditions that give rise to minimum shrinkage. Be sure that these estimated conditions are contained in the current experimental design.
  - (c) Compute the predicted value and its standard error at your recommended set of conditions.
  - (d) Does this experiment give rise to suggestions for future experimental runs? Explain.
- 6.7** Consider the reaction yield data shown in Table E6.6.
- (a) Fit a second-order model to the data.
  - (b) Construct contour plots, and comment on the fitted response surface.
  - (c) Perform the canonical analysis.

**TABLE E6.5 Data for Exercise 6.6**

| $x_1$  | $x_2$  | $x_3$  | $y$    |
|--------|--------|--------|--------|
| -1     | -1     | -1     | 0.5269 |
| +1     | -1     | -1     | 2.6680 |
| -1     | +1     | -1     | 1.0310 |
| +1     | +1     | -1     | 2.3380 |
| -1     | -1     | +1     | 3.2220 |
| +1     | -1     | +1     | 4.0060 |
| -1     | +1     | +1     | 3.3640 |
| +1     | +1     | +1     | 3.2200 |
| 0      | 0      | 0      | 2.0280 |
| 0      | 0      | 0      | 1.9670 |
| 0      | 0      | 0      | 1.9300 |
| 0      | 0      | 0      | 2.0130 |
| -1.732 | 0      | 0      | 1.6200 |
| 1.732  | 0      | 0      | 3.4510 |
| 0      | -1.732 | 0      | 2.9590 |
| 0      | 1.732  | 0      | 2.6990 |
| 0      | 0      | -1.732 | 0.9414 |
| 0      | 0      | 1.732  | 3.7880 |

**TABLE E6.6 Data for Exercise 6.7**

| $x_1$  | $x_2$  | $y$   |
|--------|--------|-------|
| -1     | -1     | 88.55 |
| -1     | 1      | 86.29 |
| 1      | -1     | 85.80 |
| 1      | 1      | 80.40 |
| 0      | 0      | 91.21 |
| 0      | 0      | 91.85 |
| -1.414 | 0      | 85.50 |
| 1.414  | 0      | 85.39 |
| 0      | -1.414 | 86.22 |
| 0      | 1.414  | 85.70 |
| 0      | 0      | 91.31 |
| 0      | 0      | 91.94 |

**6.8** In the computer industry, it is important to use experimental design and response surface methods to analyze performance data from integrated circuit/packet-switched computer networks. Here, the design variables are network input variables that affect network size, traffic load, link capacity, and so on. The response measures represent the network performance. In this experiment [Schmidt and Launsby (1990)] the design variables are:

- CS: Circuit switch arrival rate (voice calls/min)
- PS: Packet switch arrival rate (packets/sec)
- SERV: Voice call service rate (sec)
- SLOTS: Number of time slots per link (a capacity indicator)

Two responses were measured. They were:

|      |                                 |
|------|---------------------------------|
| ALU: | Average link utilization        |
| BLK: | Fraction of voice calls blocked |

The data are given in Table E6.7.

- (a) Write the design in coded form. Use the coding

$$x_1 = (\text{CS} - 4)/2$$

$$x_2 = (\text{PS} - 300)/150$$

$$x_3 = (\text{SERV} - 180)/60$$

$$x_4 = (\text{SLOTS} - 40)/6$$

- (b) Fit two second-order models for the two responses.  
 (c) Consider the response surface for the BLK response. Can the model be improved by a power transformation? Discuss.  
 (d) Use the two models above to determine optimum conditions for the two responses *separately*. We need to maximize ALU and minimize BLK.
- 6.9** In Schmidt and Launsby (1990), a simulation program was given that simulates an auto-bumper plating process using thickness as the response with time, temperature, and pH as the design variables. An experiment was conducted in order that a response surface optimization could be accomplished. The coding for the design variables is given by

$$x_1 = \frac{\text{time} - 8}{4}$$

$$x_2 = \frac{\text{temp} - 24}{8}$$

$$x_3 = \frac{\text{nickel} - 14}{4}$$

The design in the natural units is given in Table E6.8 along with the thickness values.

- (a) Write the design in coded form. Name the type of design.  
 (b) Fit a complete second-order model in the coded metric.  
 (c) Edit the model by eliminating obviously insignificant terms.  
 (d) Check model assumptions by plotting residuals against the design variables separately.  
 (e) The purpose of this experiment was not to merely find optimum conditions (conditions that maximize thickness) but to gain an impression about the role of the three design variables. Show contour plots, fixing levels of nickel at 10, 14, and 18.  
 (f) Use the plots to produce a recommended set of conditions that maximize thickness.  
 (g) Compute the standard error of prediction at the location of maximum thickness.

**TABLE E6.7 Data for Exercise 6.8**

| OBS | CS | PS  | SERV | SLOTS | ALU   | BLK   |
|-----|----|-----|------|-------|-------|-------|
| 1   | 2  | 150 | 120  | 34    | 0.282 | 0.000 |
| 2   | 2  | 150 | 120  | 46    | 0.194 | 0.000 |
| 3   | 2  | 150 | 240  | 34    | 0.372 | 0.000 |
| 4   | 2  | 150 | 240  | 46    | 0.275 | 0.000 |
| 5   | 2  | 450 | 120  | 34    | 0.553 | 0.016 |
| 6   | 2  | 450 | 120  | 46    | 0.396 | 0.000 |
| 7   | 2  | 450 | 240  | 34    | 0.676 | 0.031 |
| 8   | 2  | 450 | 240  | 46    | 0.481 | 0.003 |
| 9   | 6  | 150 | 120  | 34    | 0.520 | 0.012 |
| 10  | 6  | 150 | 120  | 46    | 0.386 | 0.000 |
| 11  | 6  | 450 | 120  | 46    | 0.594 | 0.016 |
| 12  | 6  | 300 | 180  | 40    | 0.231 | 0.000 |
| 13  | 4  | 300 | 180  | 40    | 0.325 | 0.000 |
| 14  | 4  | 300 | 180  | 52    | 0.427 | 0.000 |
| 15  | 4  | 300 | 180  | 40    | 0.561 | 0.016 |
| 16  | 4  | 300 | 180  | 40    | 0.571 | 0.005 |
| 17  | 4  | 300 | 180  | 40    | 0.561 | 0.011 |
| 18  | 4  | 300 | 180  | 40    | 0.581 | 0.020 |
| 19  | 4  | 300 | 180  | 40    | 0.562 | 0.017 |
| 20  | 4  | 300 | 180  | 40    | 0.608 | 0.036 |
| 21  | 4  | 300 | 180  | 40    | 0.560 | 0.007 |
| 22  | 5  | 400 | 210  | 43    | 0.739 | 0.070 |
| 23  | 5  | 400 | 210  | 43    | 0.746 | 0.082 |
| 24  | 5  | 400 | 150  | 37    | 0.725 | 0.067 |
| 25  | 5  | 400 | 150  | 37    | 0.725 | 0.084 |
| 26  | 5  | 400 | 150  | 43    | 0.615 | 0.027 |
| 27  | 5  | 400 | 210  | 40    | 0.785 | 0.106 |
| 28  | 5  | 400 | 210  | 40    | 0.782 | 0.125 |
| 29  | 4  | 400 | 210  | 37    | 0.747 | 0.084 |
| 30  | 4  | 400 | 210  | 37    | 0.745 | 0.101 |
| 31  | 4  | 400 | 210  | 40    | 0.695 | 0.063 |
| 32  | 4  | 400 | 210  | 40    | 0.667 | 0.057 |
| 33  | 4  | 400 | 180  | 40    | 0.648 | 0.035 |
| 34  | 5  | 400 | 180  | 40    | 0.722 | 0.074 |
| 35  | 5  | 400 | 180  | 40    | 0.731 | 0.077 |
| 36  | 5  | 400 | 180  | 37    | 0.775 | 0.110 |
| 37  | 5  | 400 | 180  | 37    | 0.779 | 0.139 |
| 38  | 4  | 400 | 180  | 37    | 0.712 | 0.057 |
| 39  | 4  | 400 | 180  | 37    | 0.701 | 0.065 |
| 40  | 3  | 400 | 210  | 37    | 0.652 | 0.017 |
| 41  | 3  | 400 | 210  | 43    | 0.541 | 0.003 |
| 42  | 3  | 400 | 150  | 37    | 0.570 | 0.003 |
| 43  | 3  | 400 | 150  | 43    | 0.476 | 0.005 |

**TABLE E6.8** Data for Exercise 6.9

| Run Number | Time | Temperature | Nickel | Thickness |
|------------|------|-------------|--------|-----------|
| 1          | 4    | 16          | 10     | 113       |
| 2          | 12   | 16          | 10     | 756       |
| 3          | 4    | 32          | 10     | 78        |
| 4          | 12   | 32          | 10     | 686       |
| 5          | 4    | 16          | 18     | 87        |
| 6          | 12   | 16          | 18     | 788       |
| 7          | 4    | 32          | 18     | 115       |
| 8          | 12   | 32          | 18     | 696       |
| 9          | 4    | 16          | 10     | 99        |
| 10         | 12   | 16          | 10     | 739       |
| 11         | 4    | 32          | 10     | 10        |
| 12         | 12   | 32          | 10     | 712       |
| 13         | 4    | 16          | 18     | 159       |
| 14         | 12   | 16          | 18     | 776       |
| 15         | 4    | 32          | 18     | 162       |
| 16         | 12   | 32          | 18     | 759       |
| 17         | 8    | 24          | 14     | 351       |
| 18         | 8    | 24          | 14     | 373       |
| 19         | 8    | 24          | 14     | 353       |
| 20         | 8    | 24          | 14     | 321       |
| 21         | 12   | 24          | 14     | 736.1     |
| 22         | 4    | 24          | 14     | 96.0      |
| 23         | 8    | 32          | 14     | 328.9     |
| 24         | 8    | 16          | 14     | 303.5     |
| 25         | 8    | 24          | 18     | 358.2     |
| 26         | 8    | 24          | 10     | 347.7     |

- 6.10** Refer to Exercise 6.9. The same simulation model was used to construct another response surface using a different experimental design. This second design, in natural units, is given in Table E6.9 along with the thickness data:
- Using the same coding as in Exercise 6.9, write out the design matrix in coded form.
  - Is this a central composite design?
  - Fit and edit a second-order response surface model. That is, fit and eliminate insignificant terms.
  - Find conditions that maximize thickness, with the constraint that the condition falls inside the design region. Use an appropriate software package. In addition, compute the standard error of prediction at the location of optimum conditions.
- 6.11** Reconsider Exercises 6.9 and 6.10. In addition to the designs in these exercises, a *D-optimal* design was used to generate data on thickness from the simulation model. The notion of ***D-optimality*** was introduced in Chapter 3 and will be discussed in more depth in Chapter 9. The design and data are given in Table E6.10.
- Using the same coding scheme as in Exercises 6.9 and 6.10, write the design in coded units.

**TABLE E6.9** Data for Exercise 6.10

| Run Number | Time | Temperature | Nickel | Thickness |
|------------|------|-------------|--------|-----------|
| 1          | 4    | 16          | 14     | 122       |
| 2          | 12   | 16          | 14     | 790       |
| 3          | 4    | 32          | 14     | 100       |
| 4          | 12   | 32          | 14     | 695       |
| 5          | 4    | 24          | 10     | 50        |
| 6          | 12   | 24          | 10     | 720       |
| 7          | 4    | 24          | 18     | 118       |
| 8          | 12   | 24          | 18     | 650       |
| 9          | 8    | 16          | 10     | 330       |
| 10         | 8    | 32          | 10     | 314       |
| 11         | 8    | 16          | 18     | 302       |
| 12         | 8    | 32          | 18     | 340       |
| 13         | 8    | 24          | 14     | 364       |
| 14         | 8    | 24          | 14     | 342       |
| 15         | 8    | 24          | 14     | 404       |

**TABLE E6.10** Data for Exercise 6.11

| Run Number | Time | Temperature | Nickel | Thickness |
|------------|------|-------------|--------|-----------|
| 1          | 12   | 32          | 10     | 717       |
| 2          | 4    | 16          | 10     | 136       |
| 3          | 12   | 16          | 18     | 787       |
| 4          | 4    | 32          | 18     | 78        |
| 5          | 4    | 16          | 18     | 87        |
| 6          | 4    | 32          | 10     | 80        |
| 7          | 12   | 16          | 10     | 760       |
| 8          | 8    | 24          | 10     | 282       |
| 9          | 8    | 32          | 14     | 318       |
| 10         | 4    | 24          | 14     | 96        |
| 11         | 12   | 32          | 18     | 682       |
| 12         | 12   | 24          | 14     | 747       |
| 13         | 12   | 24          | 14     | 764       |
| 14         | 12   | 24          | 14     | 742       |

- (b) Fit a second-order model, and edit it—that is, eliminate insignificant terms.
- (c) Find conditions of estimated maximum thickness. Can this be done with a two-dimensional contour plot? Explain. Compute the standard error of prediction at the set of optimum conditions.

**6.12** Exercises 6.9, 6.10, and 6.11 illustrate the use of three different types of designs for modeling the same system. All three designs are used because of their capability to efficiently fit a *complete* second-order model, even if the edited model contains no interactions. Another type of design that allows for estimation of linear and pure quadratic effects is the **Taguchi**  $L_{27}$ . The Taguchi approach to quality improvement will be discussed in detail in Chapter 11. The  $L_{27}$  was also used to generate data

TABLE E6.11 Data for Exercise 6.12

| Run Number | Time | Temperature | Nickel | Thickness |
|------------|------|-------------|--------|-----------|
| 1          | 4    | 16          | 10     | 113       |
| 2          | 4    | 16          | 14     | 152       |
| 3          | 4    | 16          | 18     | 147       |
| 4          | 4    | 24          | 10     | 65        |
| 5          | 4    | 24          | 14     | 53        |
| 6          | 4    | 24          | 18     | 130       |
| 7          | 4    | 32          | 10     | 83        |
| 8          | 4    | 32          | 14     | 40        |
| 9          | 4    | 32          | 18     | 94        |
| 10         | 8    | 16          | 10     | 339       |
| 11         | 8    | 16          | 14     | 303       |
| 12         | 8    | 16          | 18     | 381       |
| 13         | 8    | 24          | 10     | 348       |
| 14         | 8    | 24          | 14     | 342       |
| 15         | 8    | 24          | 18     | 420       |
| 16         | 8    | 32          | 10     | 327       |
| 17         | 8    | 32          | 14     | 255       |
| 18         | 8    | 32          | 18     | 322       |
| 19         | 12   | 16          | 10     | 745       |
| 20         | 12   | 16          | 14     | 780       |
| 21         | 12   | 16          | 18     | 740       |
| 22         | 12   | 24          | 10     | 772       |
| 23         | 12   | 24          | 14     | 769       |
| 24         | 12   | 24          | 18     | 755       |
| 25         | 12   | 32          | 10     | 735       |
| 26         | 12   | 32          | 14     | 726       |
| 27         | 12   | 32          | 18     | 757       |

on thickness as in the previous exercises. The design and response data are given in Table E6.11.

- (a) Code the design factors as in previous exercises.
- (b) Form the complete matrix  $\mathbf{X}$  for the  $L_{27}$  (using coded design levels).
- (c) Comment on whether or not this design is adequate for a *complete* second-order model (including interactions).

- 6.13** In this exercise we will make comparisons among the central composite, Box–Behnken, and  $D$ -optimal designs used in Exercises 6.9, 6.10, and 6.11. We will not deal with any particular edited model, but rather with a complete second-order model. Compute the *scaled* standard error of prediction

$$\sqrt{\frac{N \operatorname{Var}[\hat{y}(\mathbf{x})]}{\sigma^2}} = \sqrt{N[\mathbf{x}^{(2)'}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}^{(2)}]}$$

at the locations given in Table E6.12 in coded design units.  $N$  is the size of the design. Comment on comparisons among the three designs.

- 6.14** The experiment given in Table E6.13 uses a central composite design to study three

**TABLE E6.12 Design Location for Exercise 6.13**

| $x_1$ | $x_2$ | $x_3$ |
|-------|-------|-------|
| -1    | -1    | -1    |
| 1     | -1    | -1    |
| -1    | 1     | -1    |
| -1    | -1    | 1     |
| 1     | 1     | -1    |
| 1     | -1    | 1     |
| -1    | 1     | 1     |
| 1     | 1     | 1     |
| -1    | 0     | 0     |
| 1     | 0     | 0     |
| 0     | -1    | 0     |
| 0     | 1     | 0     |
| 0     | 0     | -1    |
| 0     | 0     | 1     |
| 0     | 0     | 0     |

variables in a chemical process.

- Fit a second-order model to the conversion response  $y_1$ .
- Investigate the fitted surface using contour plots.
- Find the stationary point, and perform the canonical analysis.

**6.15** Consider the viscosity response data in Table 6.8.

**TABLE E6.13 Data for Exercise 6.14**

| $x_1$ Time | $x_2$ Temperature | $x_3$ Catalyst | $y_1$ Conversion | $y_2$ Activity |
|------------|-------------------|----------------|------------------|----------------|
| -1         | -1                | -1             | 74.00            | 53.20          |
| 1          | -1                | -1             | 51.00            | 62.90          |
| -1         | 1                 | -1             | 88.00            | 53.40          |
| 1          | 1                 | -1             | 70.00            | 62.60          |
| -1         | -1                | 1              | 71.00            | 57.30          |
| 1          | -1                | 1              | 90.00            | 67.90          |
| -1         | 1                 | 1              | 66.00            | 59.80          |
| 1          | 1                 | 1              | 97.00            | 67.80          |
| -1.682     | 0                 | 0              | 76.00            | 59.10          |
| 1.682      | 0                 | 0              | 79.00            | 65.90          |
| 0          | -1.682            | 0              | 85.00            | 60.00          |
| 0          | 1.682             | 0              | 97.00            | 60.70          |
| 0          | 0                 | 1.682          | 55.00            | 57.40          |
| 0          | 0                 | 1.682          | 81.00            | 63.20          |
| 0          | 0                 | 0              | 81.00            | 59.20          |
| 0          | 0                 | 0              | 75.00            | 60.40          |
| 0          | 0                 | 0              | 76.00            | 59.10          |
| 0          | 0                 | 0              | 83.00            | 60.60          |
| 0          | 0                 | 0              | 80.00            | 60.80          |
| 0          | 0                 | 0              | 91.00            | 58.90          |



- (a) Perform a canonical analysis on the second-order model for these data.  
 (b) Use the double linear regression method to find confidence intervals on the eigenvalues. The design in terms of the canonical variables is

$$\mathbf{U} = \begin{bmatrix} 0.8922 & 1.0972 \\ -1.0972 & 0.8922 \\ 1.0972 & -0.8922 \\ -0.8922 & -1.0972 \\ 1.4066 & 0.1449 \\ -1.4066 & -0.1449 \\ -0.1449 & 1.4066 \\ 0.1449 & -1.4066 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

- (c) Interpret the fitted surface. Can you suggest a procedure for making the fitted model more closely approximate the true response surface?

**6.16** For the estimated response surface listed below, answer the following questions:

$$\hat{y} = 35 + 1.02x_1 + 1.76x_2 - 0.59x_3 - 2.66x_1x_2 + 3.02x_1x_3 - 0.62x_2x_3 + 5.04x_1^2 + 2.63x_2^2 + 5.33x_3^2$$

- (a) Write the model in the canonical form.  
 (b) Describe what kind of stationary point the response surface has and where it is located.  
 (c) If all the coded variables have range  $[-1, 1]$  and the range of the actual variables are as follows:  $X_1^* \in [10, 30]$ ,  $X_2^* \in [150, 200]$ ,  $X_3^* \in [1.0, 2.0]$   
 (i) Where is the global minimum located in terms of the original variables?  
 (ii) Where is the maximum located within the observed range of the actual variables (stay in the cube design space)?

**6.17** For a given problem we obtain a second-order model, which we describe in terms of its canonical parameterization. We obtain the following pieces for the model in the form  $\hat{y} = 10.5 + \mathbf{xb} + \mathbf{x\hat{B}x}$

$$\mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \hat{\mathbf{B}} = \begin{pmatrix} 2.04 & -0.1 & -0.075 \\ -0.1 & 0.160 & 0.25 \\ -0.075 & 0.25 & 1.15 \end{pmatrix}, \quad \mathbf{x}_s = \begin{pmatrix} -0.23 \\ 0.85 \\ -0.63 \end{pmatrix}$$

with eigenvalues and eigenvectors:

|                    |             |            |             |
|--------------------|-------------|------------|-------------|
| <b>\$ values:</b>  |             |            |             |
| [1]                | 2.05420808  | 1.19864128 | 0.09715064  |
| <b>\$ vectors:</b> |             |            |             |
|                    | [, 1]       | [, 2]      | [, 3]       |
| [1, ]              | 0.99276672  | -0.1127446 | -0.04126621 |
| [2, ]              | -0.06567526 | -0.2222498 | -0.97277530 |
| [3, ]              | -0.10050377 | -0.9684491 | 0.22804674  |

- (a) Write the model in standard form ( $y = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \dots$ ) with the actual coefficients listed for all main, interaction, and quadratic effects.
- (b) Describe the kind of surface we have for the response.
- (c) If we are using a standard cuboidal region, is the stationary point located inside the region of study? Explain.

If the value of the function at the stationary point is 11.91, give the equation for the model in the canonical form,  $\hat{y} = \hat{y}_s + \sum \lambda_i w_i$ .

- 6.18** Martins *et al.* (2009) investigate the effect of several factors on a microparticle-based DNA/DNA hybridization assay. The measured response is the four factors considered are immobilization area ( $X_1$ ), probe concentration ( $X_2$ ), hybridization area ( $X_3$ ) and target concentration ( $X_4$ ). A central composite design was performed with a total of 27 runs. The data are shown with the factors already scaled in Table. E6.14.
- (a) Fit a second-order model to the response, fluorescence. What fraction of the variability in fluorescence is explained by the full model.
  - (b) Find a reduced model by removing nonsignificant terms. What fraction of the variability is explained by the reduced model.
  - (c) Draw contour plots of the response as a function of each pair of variables ( $X_1$  &  $X_2$ ,  $X_1$  &  $X_3$ , ...) with the other variables held fixed at their center values. Which variables are most influential in changing the values of the response?
  - (d) Where in the design region does the model predict the maximum fluorescence value? How does this value compare to the observed value at (or closest to) this location?
- 6.19** Consider the data described in Exercise 6.18 and given in Table E6.14. The following describe the factor levels of  $X_1$ – $X_4$  in the original units.

| Level | Immobilization Area, $\xi_1$ (cm <sup>2</sup> ) | Probe Concentration, $\xi_2$ ( $\mu$ M) | Hybridization Area, $\xi_3$ (cm <sup>2</sup> ) | Target Concentration, $\xi_4$ ( $\mu$ M) |
|-------|---|---|--|--|
| -1    | 0.35  | 0.025                                   | 0.35   | 0.025                                    |
| 0     | 3.5   | 2.5                                     | 2.0  | 2.5                                      |
| +1    | 7.0   | 5.0                                     | 3.5  | 5.0                                      |

TABLE E6.14 Data for Exercises 6.18 and 6.19

| $X_1$ | $X_2$ | $X_3$ | $X_4$ | Fluorescence |
|-------|-------|-------|-------|--------------|
| -1    | -1    | -1    | -1    | 4.7          |
| 1     | -1    | -1    | -1    | 4.7          |
| -1    | 1     | -1    | -1    | 28           |
| 1     | 1     | -1    | -1    | 81.2         |
| -1    | -1    | 1     | -1    | 5.7          |
| 1     | -1    | 1     | -1    | 3.8          |
| -1    | 1     | 1     | -1    | 12.2         |
| 1     | 1     | 1     | -1    | 19.5         |
| -1    | -1    | -1    | 1     | 4.4          |
| 1     | -1    | -1    | 1     | 2.6          |
| -1    | 1     | -1    | 1     | 83.7         |
| 1     | 1     | -1    | 1     | 84.7         |
| -1    | -1    | 1     | 1     | 6.8          |
| 1     | -1    | 1     | 1     | 2.4          |
| -1    | 1     | 1     | 1     | 76           |
| 1     | 1     | 1     | 1     | 77.9         |
| -1    | 0     | 0     | 0     | 42.6         |
| 1     | 0     | 0     | 0     | 52.3         |
| 0     | -1    | 0     | 0     | 2.6          |
| 0     | -1    | 0     | 0     | 72.8         |
| 0     | 0     | -1    | 0     | 47.7         |
| 0     | 0     | 1     | 0     | 54.4         |
| 0     | 0     | 0     | -1    | 30.8         |
| 0     | 0     | 0     | 1     | 64.8         |
| 0     | 0     | 0     | 0     | 51.6         |
| 0     | 0     | 0     | 0     | 52.6         |
| 0     | 0     | 0     | 0     | 56.1         |

- (a) In Exercise 6.18, the authors of the study reported that the factor levels are equally spaced  $(-1, 0, +1)$  for all factors. However, when the actual levels of the experiment are shown, there are some small discrepancies in the spacing. By matching the factor levels for 0, and +1, what are the actual scaled factor levels for each of the factors?
- (b) Reanalyze the data with the new scaled factor levels. Compare the full model and the reduced models for the analysis in Exercise 6.18 with what is obtained. Comment on how this might affect conclusions.
- (c) Using the new analysis, compute the standard error of prediction at the following locations:
- (i)  $\xi_1 = 3.5 \text{ cm}^2$ ,  $\xi_2 = 2.5 \text{ }\mu\text{M}$ ,  $\xi_3 = 2.0 \text{ cm}^2$ ,  $\xi_4 = 2.5 \text{ }\mu\text{M}$
  - (ii)  $\xi_1 = 7.0 \text{ cm}^2$ ,  $\xi_2 = 5.0 \text{ }\mu\text{M}$ ,  $\xi_3 = 3.5 \text{ cm}^2$ ,  $\xi_4 = 5.0 \text{ }\mu\text{M}$
  - (iii)  $\xi_1 = 0.35 \text{ cm}^2$ ,  $\xi_2 = 0.025 \text{ }\mu\text{M}$ ,  $\xi_3 = 0.35 \text{ cm}^2$ ,  $\xi_4 = 0.025 \text{ }\mu\text{M}$
  - (iv)  $\xi_1 = 3.5 \text{ cm}^2$ ,  $\xi_2 = 2.5 \text{ }\mu\text{M}$ ,  $\xi_3 = 0.35 \text{ cm}^2$ ,  $\xi_4 = 2.5 \text{ }\mu\text{M}$
  - (v)  $\xi_1 = 3.5 \text{ cm}^2$ ,  $\xi_2 = 2.5 \text{ }\mu\text{M}$ ,  $\xi_3 = 3.5 \text{ cm}^2$ ,  $\xi_4 = 2.5 \text{ }\mu\text{M}$

- (d) By comparing the results of (c) ii and iii, comment on the impact of the non-standard scaling on the standard error of prediction at the factorial locations.
- (e) By comparing the results of (c) iv and v, comment on the impact of the nonstandard scaling on the standard error of prediction at the axial locations for factor  $X_3$ .
- (f) Why might the experimenters have not used equal spacing? (i.e., why did they not set  $X_i = 0$ ?) Comment on the trade-offs between using a standard design with fixed levels, and adapting the design to have unequal spacing.

**6.20** A five level central composite design was used in an experiment to study the ability of different treatments to inhibit the growth of *Listeria monocytogenes* strains. The results for the strain, ATCC 19114, are given in Table E6.15.

- (a) Convert the factor levels to scaled factors, using first 8 observations to define a 24 factorial with scaled ranges  $[-1, +1]$ .
- (b) Fit a second-order model to the observed data, and find a suitable reduced model that describes how the different factors influence the response.
- (c) If the goal is to minimize growth, what location in the design space is best to achieve this goal?
- (d) What is the predicted growth amount at the location identified in (c)? Give 95% confidence interval and 95% prediction intervals for the response at this location?
- (e) A control run, which matches conditions where no preventative measures were taken to inhibit growth, was also performed with the following levels:

| Run | NaCl | pH | MaxExtraction | Growth |
|-----|------|----|---------------|--------|
| 18  | 0    | 7  | 0             | 0.70   |

What does the model predict for the growth amount at this location? How well do you think that the model fits here? (Hint: compare a 95% prediction interval to the observation)

- 6.21** Consider the data described in Exercise 6.20 and given in Table E6.15.
- (a) Find the region of the design space where growth for the strain, ATCC 19114, is estimated to be greater than 0.5.
  - (b) If we wish to be cautious about where we might find growth values above 0.5, we might examine the upper bound of an uncertainty region. Find the region of the design space where the upper bound of a 95% confidence interval for growth is estimated to be greater than 0.5.
  - (c) Find the region of the design space where the upper bound of a 95% prediction interval for growth is predicted to be greater than 0.5.
  - (d) By comparing the results of (a), (b), and (c), discuss the role of model uncertainty and natural variability (the variance of  $\sigma$ ) on changing the size of the region where we may be concerned about large values of growth.
  - (e) Describe a situation where the results from each of (a), (b), and (c) might be used to interpret the results of the experiment.

**TABLE E6.15** Data for Exercises 6.20, 6.21, and 6.22

| Run | NaCl | pH  | MaxExtraction | Growth |
|-----|------|-----|---------------|--------|
| 1   | 1.5  | 6.5 | 0.156         | 0.48   |
| 2   | 0.5  | 6.5 | 0.156         | 0.52   |
| 3   | 1.5  | 6.5 | 0.078         | 0.61   |
| 4   | 0.5  | 6.5 | 0.078         | 0.62   |
| 5   | 1.5  | 5.5 | 0.156         | 0.29   |
| 6   | 0.5  | 5.5 | 0.156         | 0.30   |
| 7   | 1.5  | 5.5 | 0.078         | 0.35   |
| 8   | 0.5  | 5.5 | 0.078         | 0.40   |
| 9   | 1    | 6   | 0.117         | 0.35   |
| 10  | 1    | 7   | 0.117         | 0.58   |
| 11  | 1    | 5   | 0.117         | 0.27   |
| 12  | 1    | 6   | 0.195         | 0.00   |
| 13  | 1    | 6   | 0.039         | 0.58   |
| 14  | 2    | 6   | 0.117         | 0.32   |
| 15  | 0    | 6   | 0.117         | 0.42   |
| 16  | 1    | 6   | 0.117         | 0.37   |
| 17  | 1    | 6   | 0.117         | 0.37   |

- 6.22** Consider the data described in Exercise 6.20 and given in Table E6.15.
- If the experimenter has chosen to fit a first-order model with all of the two-way interactions, find the best reduced model based on this initial assumption.
  - Find the standard error of prediction for the best model based on a second-order model and the best model for a first-order model with interactions at the following locations:
    - the center of the design space
    - the axial points
    - the factorial points
  - Based on the model found in (a), find the region of the design space where growth for the strain, ATCC 19114, is estimated to be greater than 0.5.
  - By comparing the results from (c) to Exercise 6.21(a) results, discuss the trade-off between the choice of the two models for accurate and precise estimation.

### Short Answer Questions

- 6.23** Answer the following short answer questions:
- Justify why using a low-order polynomial is a helpful strategy when modeling many relationships between input factors and responses.
  - Describe a situation where a stationary point is beneficial for optimizing the response.
  - Describe a situation where a stationary point is not that helpful for optimizing the response.
  - When can a ridge system be advantageous when optimizing a response?

- (e) Describe how the standard error of the predicted response is connected to the chosen designed experiment.
- (f) Describe the connection between the standard error of the predicted response and the location of the stationary point.

**6.24** Answer the following True/False questions:

- (a) Using a second-order model to estimate the underlying relationship is a common approach to approximating the surface.
- (b) The location of the stationary point corresponds to a global maximum or minimum of the surface.
- (c) Examining the canonical variables provides information about the nature of the surface.
- (d) Combining model reduction or model selection with the use of canonical variables provides flexibility in interpreting the shape of the response.
- (e) Uncertainty in the model parameters can influence the nature of the response through the values of the eigenvalues.
- (f) The existence of a ridge system in the canonical variables can restrict the choice of optimal location.
- (g) Whether the location of the stationary point is inside or outside of the design region changes the usefulness of the canonical analysis.
- (h) The magnitude of the natural variability of the response influences the size of the confidence region for location of the stationary point.
- (i) If the sign of all of the eigenvalues is the same, the response surface will have a global maximum.

---

# 7

---

## MULTIPLE RESPONSE OPTIMIZATION

### 7.1 BALANCING MULTIPLE OBJECTIVES

Up to this point we have emphasized the use of response surface analyses for improvement of the quality of products and processes. We have focused on modeling a measured response or a function of design variables and letting the analysis indicate areas in the design region where the process is likely to give desirable results, the term “desirable” being a function of the predicted response. However, in many instances the term “desirable” is a function of more than one response. Intuitively it should not be surprising that when an experimenter dedicates the time and effort to run a thoughtfully designed experiment, that it often makes sense to gather information on multiple characteristics that result from the manipulation of the inputs. In addition, the quality or desirability of a product or process cannot be simplified to a goodness of a single attribute.

In a chemical process there are almost always several properties of the product output that must be considered in the definition of “desirable.” In many consumer products (food or beverages), the scientist must deal with taste as a response but also consider other responses such as color and texture, as well as undesirable by-products. In the pharmaceutical or biomedical area, the clinician is primarily concerned with the efficacy of the drug or remedy but must not ignore the possibility of serious side effects. In a tool life problem, cutting speed ( $x_1$ ) and depth of cut ( $x_2$ ) influence the primary response, the life of the tool. However, a secondary response, rate of metal removed, may also be important in the study.

As precious resources are stretched thinner and industry seeks to extract more information from a single experiment, having sound strategies for examining and combining multiple objectives is increasingly important. Before considering how to summarize and optimize several responses from a designed experiment, we first discuss a more general framework for examining the relationship between several competing goals.

There are several aspects to this problem that complicate the process of choosing a best alternative when considering multiple facets to the decision-making. First, there are almost no practical decisions where it is possible to get the optimal result for each criterion individually in a single choice. In the rare cases where this might occur, no special methods are needed to select the best option. Therefore, for most situations, it is necessary to make trade-offs between the different objectives: Does it make sense to sacrifice a bit of this to get a bit more of that?

Secondly, many of the characteristics which are being considered in the decision are measured on different scales: How do we trade-off some additional cost in exchange for an improvement in the quality. Cost might be measured in dollars or time, while quality is measured on a completely different scale.

Thirdly, how these different characteristics are valued may differ depending on the decision-maker. There is no global consensus on what trade-off might be ideal between cost and quality, but rather many right answers for different sets of priorities. So this part of the process inherently has a subjective component where individual priorities are integrated into the process. For a low-end manufacturer, reducing quality to save on production costs might be the ideal choice. For a higher-end producer, the risk of a dissatisfied customer may trump a fractional discount on production costs.

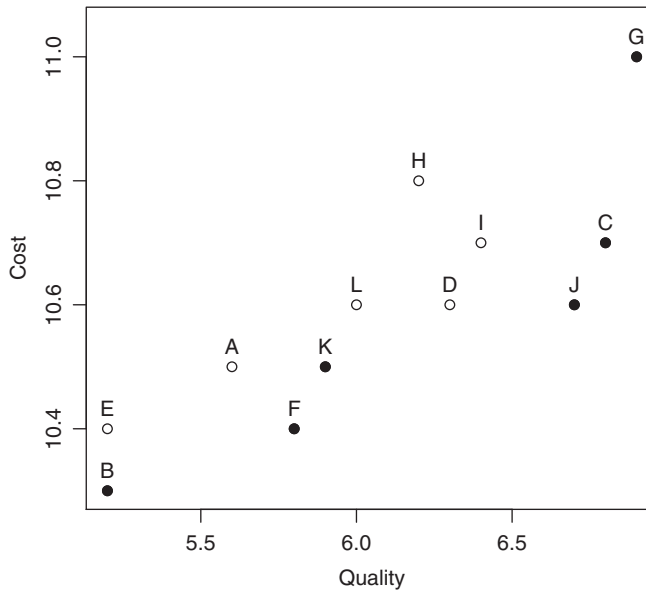
As a result, methods for evaluating and optimizing based on multiple objectives should provide mechanisms for incorporating the experimenter's priorities and preferences.

**Example 7.1 Choosing a Supplier** In a manufacturing application, there are several potential suppliers of a particular raw material, each with associated costs and quality. The goal of selecting a supplier is to balance the trade-offs between these characteristics to find a suitable source for the raw material. Table 7.1 shows the set of possible manufacturers with their associated costs and measured quality of product. Examining the scatterplot of alternatives in Fig. 7.1, the ideal is to be located in the bottom right corner of the plot, where cost is minimized and quality is maximized. As is typically the case, no supplier simultaneously offers this combination, but rather the decision-maker will need to make trade-offs between cost and quality. However, it is also clear that some of the options do not make sense. For example, it would not be rational to select supplier D (cost = 10.6 and

**TABLE 7.1 Supplier Cost and Quality**

| Supplier | Cost | Quality |
|----------|------|---------|
| A        | 10.5 | 5.6     |
| B        | 10.3 | 5.2     |
| C        | 10.7 | 6.8     |
| D        | 10.6 | 6.3     |
| E        | 10.4 | 5.2     |
| F        | 10.4 | 5.8     |
| G        | 11.0 | 6.9     |
| H        | 10.8 | 6.2     |
| I        | 10.7 | 6.4     |
| J        | 10.6 | 6.7     |
| K        | 10.5 | 5.9     |
| L        | 10.6 | 6.0     |





**Figure 7.1** Scatterplot of cost and quality measures for suppliers in Table 7.1. Solid circles indicate suppliers on the Pareto front.

quality = 7.3) as a preferred choice since supplier J has the same cost but higher quality (7.7). Similarly, supplier I would not be a reasonable choice since supplier C has the same cost (10.7) and better quality (7.8 instead of 7.4), or supplier J has both lower cost and better quality.

Hence only those choices shown in Fig. 7.1 with solid circles (●) do not have a superior choice among the available list of suppliers. These options are the ones from which the decision-maker should select, and they are known as the **Pareto front**. See Lu et al. (2011) for more details on the construction and interpretation of Pareto fronts. The suppliers with open circles (○) are **inferior**, since there is at least one choice on the Pareto front that is at least as good on all of the criteria, and strictly better on at least one. Therefore, a good objective first step in selecting a best supplier is to eliminate the noncontenders, or those not found on the Pareto front. With the remaining six suppliers, B, F, K, J, C, G—ordered from lowest (best) to highest (worst) cost and from lowest (worst) to highest (best) quality—we can clearly see the trade-off between the two objectives. If the lowest cost is the top priority, then supplier B is best. Alternately, if highest quality is paramount, then supplier G is best. Suppliers F, K, J, and C offer options which balance cost and quality. In JMP, the Pareto front for any number of numeric criteria can be found in the “Rows” option of a data table, under “Row Selection” and “Select Dominant.” There are then options to select which columns to consider and whether optimization requires maximizing or minimizing the quantity.

Note that choosing the cheapest supplier seems to have a disproportionate penalty with respect to quality: Dropping the price from 10.4 to 10.3 requires a drop in quality from 5.8 to 5.2. Similarly, to obtain the best quality of 6.9 (versus 6.8) has a large associated increase in price (11.0 versus 10.7). This phenomenon is quite common across many applications,

and it reflects that to optimize a single objective often requires considerable sacrifice of other criteria. Taking a small step back from the optimal value can often achieve more robust performance on other secondary objectives.

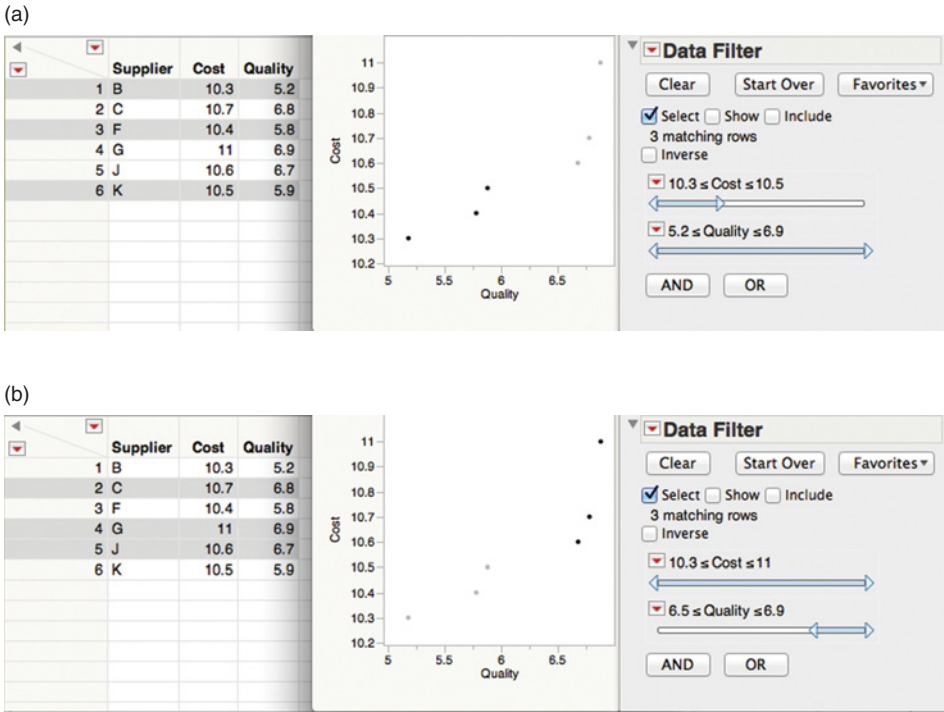
Once the Pareto front has been identified—here with the six suppliers—then which of these options is best is a subjective choice based on how cost and quality are prioritized. For one company, emphasizing cost more than quality might suggest supplier F as best, while for another company, the higher quality of supplier J or C may warrant the additional expense of a higher cost. Regardless of which of these final choices are made, the first step of identifying the Pareto front eliminates the noncontenders and simplifies the number of alternative under consideration. In addition, the Pareto front highlights the range of sensible values for each of cost (10.3 to 11.0) and (5.2 to 6.9), which provides a frame of reference to consider comparisons.

In addition, by examining the ranges of each of the objectives, we are able to assess whether these represent similar or different ranges of preference. For example, the highest cost is 6.8% larger than the smallest on the Pareto front ( $11.0/10.3 = 1.068$ ), while the highest quality is 32.7% larger than the low end of the range from the Pareto front ( $6.9/5.2 = 1.327$ ). Hence moving from best to worst on the quality scale may represent more serious changes than the change in cost. However, the percentage change does not tell the whole story, as the decision-maker should also consider whether the range covers choices that represent “excellent” to “poor,” or “very good” to “good.” This can change how much we emphasize trade-offs between alternatives.

The definition of the Pareto front based on all non-inferior candidates extends to higher dimensions with more than two criteria to be considered. For choosing between alternatives, it provides an objective straightforward method of simplifying the set of candidates to only those rational options. In higher dimensions, visualizing the Pareto front becomes more challenging, but its advantages remain unchanged.

Since for most decisions, a single best option needs to be identified, some additional tools are needed to formalize this part of the process. Which of the choices on the Pareto front makes the most sense for a particular problem depends on how the different objectives are valued, and hence involves some additional subjective input from the decision-maker. One way is to use the **threshold approach**, which tackles the individual criteria one-by-one based on their relative importance. Specifically, it may start with the most important criterion, which determines the threshold below (or above) which we are unwilling to accept values. This eliminates a subset of the choices. The process is then repeated with the next most important criterion until all criteria are examined. The advantage of this approach is being able to work with the raw data on their original scales. In JMP, the “Data Filter” allows flexible specification of thresholds with visualization of the subsets of the data. Figure 7.2 show two versions of this process. In Fig. 7.2a the decision-maker has prioritized cost and has decided to find the best option with a price at or below 10.5. Once this threshold has been determined, then selecting the supplier with the highest available quality leads to supplier K. Alternately, setting an initial threshold of quality of at least 6.5 in Fig. 7.2b leads to the choice of supplier J.

Another popular choice for selecting between alternatives is to use the simultaneous optimization technique popularized by Derringer and Suich (1980). Their procedure makes use of **desirability functions**. The thinking behind this approach is that different criteria are often measured on different scales, and a primary difficulty in knowing how to compare them comes from the “apples to oranges” incompatibility of these attributes. The desirability function suggests that we make different objectives comparable by assigning them



**Figure 7.2** Threshold approach using the data filter in JMP for (a) cost constrained to be lower than 10.5 and (b) quality constrained to be above 6.5.

a desirability score. Hence, we first convert each objective into an individual desirability function  $d_i$  that varies over the range

$$0 \leq d_i \leq 1$$

where if the objective is at its goal or target, then  $d_i = 1$ , and if the response is outside an acceptable region,  $d_i = 0$ . In this way “best” for each criterion is scored the same way (with a score of 1), and unacceptable is also comparable (with a score of 0). Then the optimal combination is chosen to maximize the overall desirability, where additive or multiplicative forms of the combined desirability score are possible. For  $m$  objectives, the additive form

$$D = \sum_{i=1}^m w_i d_i \quad (7.1)$$

allows for superior performance on one or more objectives to compensate for poor performance on another. The multiplicative form

$$D = d_1^{w_1} d_2^{w_2} \dots d_m^{w_m} \quad (7.2)$$

places additional penalties on poor performance of a single objective which is difficult to overcome, regardless of how other objectives perform. The weights  $w_i$ , which are chosen

**TABLE 7.2 Desirability Scores for Cost and Quality for Six Suppliers on Pareto Front**

| Supplier | Cost | Scaled Cost | Quality | Scaled Quality |
|----------|------|-------------|---------|----------------|
| B        | 10.3 | 1.0         | 5.2     | 0.0            |
| F        | 10.4 | 0.857       | 5.8     | 0.353          |
| K        | 10.5 | 0.714       | 5.9     | 0.412          |
| J        | 10.6 | 0.571       | 6.7     | 0.882          |
| C        | 10.7 | 0.429       | 6.8     | 0.941          |
| G        | 11.0 | 0.0         | 6.9     | 1.0            |

to sum to 1, allow different emphases to be placed on the different objectives to match the priorities of the decision.

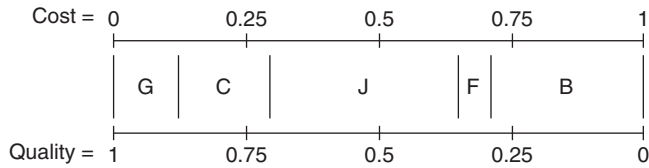
For the cost-quality example, if we scale both objectives between 0 and 1 based on the values from the Pareto front. Table 7.2, shows the 6 suppliers on the Pareto front, sorted from best to worst quality (and hence worst to best cost). By using a linear transformation, we can obtain the desirability scores for each objective for each supplier. Note how small values of cost and high values of quality have larger desirability scores to reflect the goals of the optimization.

Depending on how the decision-maker values the different criteria, different suppliers are identified as best. Table 7.3 shows several different choices of weights using the additive form of the desirability function, and the best supplier for each combination. Note how the final column matches the scaled cost column in Table 7.2, since all of the weight is placed on the cost. For the first three columns, there is some balancing of objectives being done, with a larger weight on one objective giving a result closer to that end of the Pareto front.

The idea of synthesized efficiency [Lu and Anderson-Cook (2012)] allows comparison between any choice and the best available choice at a specified weight combination. For example, if we are interested in comparing how well different alternatives compared when  $(w_C, w_Q) = (0.25, 0.75)$ , then the synthesized efficiency compares the desirability function value for a supplier to the best supplier, here supplier C with a desirability function value of 0.8130. Hence the synthesized efficiency score for suppliers B and J are  $0.25/0.8130 = 0.308$  and  $0.8043/0.8130 = 0.989$ , respectively. Hence, we can see that while supplier J is not optimal for this weight combination, it performs extremely well compared to the best available, and would represent only a small sacrifice from optimal. Obviously, the best value for a given weight combination has synthesized efficiency of value 1.

**TABLE 7.3 Additive Desirability Function Values for Suppliers for Different Weights with Best Choice for Each Weight Combination Shown in Bold**

| Supplier | $w_C = 0.25$  | $w_C = 0.5$   | $w_C = 0.75$  | $w_C = 1.0$  |
|----------|---------------|---------------|---------------|--------------|
|          | $w_Q = 0.75$  | $w_Q = 0.5$   | $w_Q = 0.25$  | $w_Q = 0.0$  |
| B        | 0.2500        | 0.5000        | <b>0.7500</b> | <b>1.000</b> |
| F        | 0.4790        | 0.6050        | 0.7310        | 0.857        |
| K        | 0.4875        | 0.5630        | 0.6385        | 0.714        |
| J        | 0.8043        | <b>0.7265</b> | 0.6488        | 0.571        |
| C        | <b>0.8130</b> | 0.6850        | 0.5570        | 0.429        |
| G        | 0.7500        | 0.5000        | 0.2500        | 0.000        |

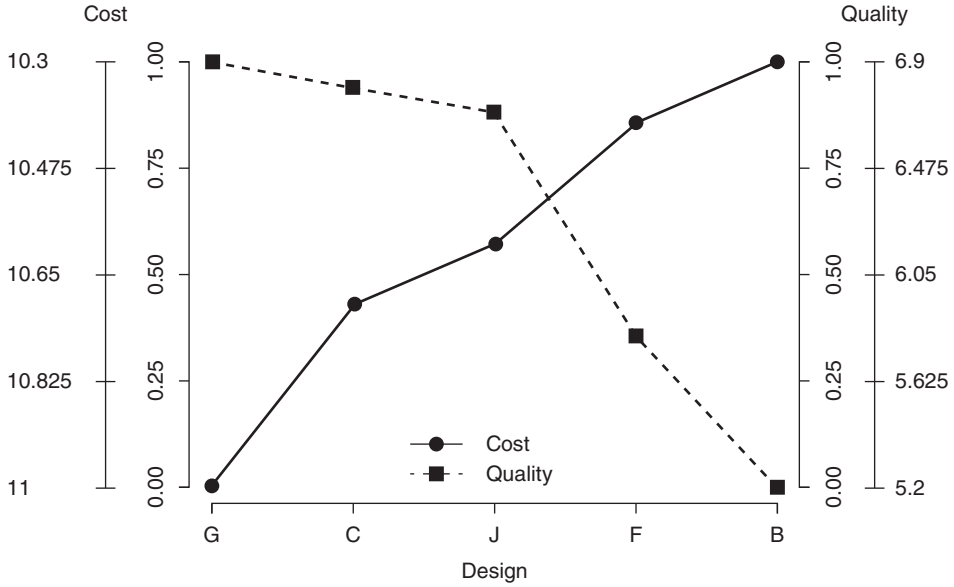


**Figure 7.3** Best supplier for each combination of desirability function weights between 0 and 1 for cost and quality.

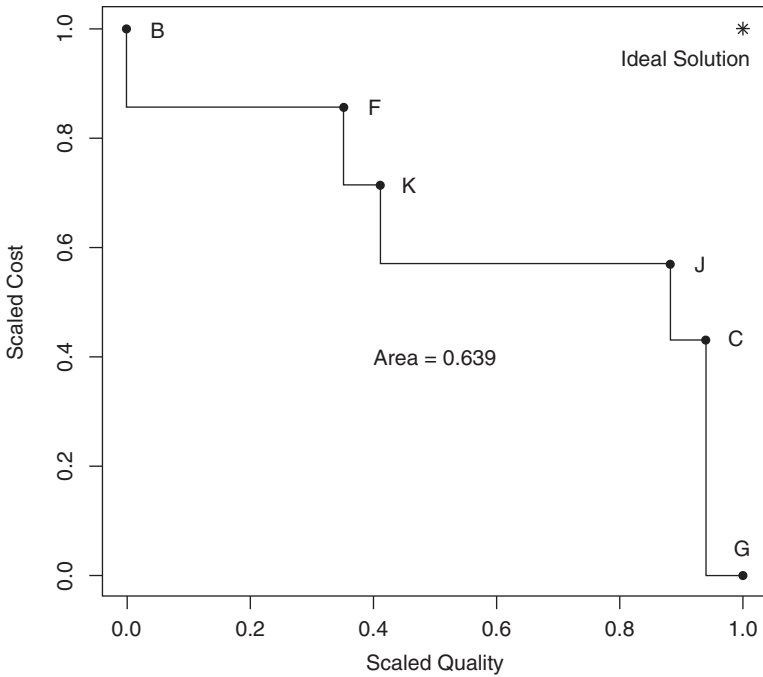
Since there may be multiple decision-makers involved in choosing a supplier, each with potentially different agendas and priorities, it is helpful to consider the robustness of a choice of supplier to different weight combinations. Figure 7.3 shows which suppliers are best for any weight combination. This plot can be easily calculated by examining a finely gridded set of weights, calculating the desirability function (with either the additive or multiplicative form) for all choices on the Pareto front, and then selecting the maximum value for each set of weights [see Lu, Anderson-Cook, and Robinson (2011)]. Note that five of the six suppliers on the Pareto front are best for at least one combination of weights, with only supplier K omitted. This often occurs for choices that are recessed from the convex hull of the Pareto front. Both suppliers B and J are best for large proportions of the weight ranges, but which one is preferred would be determined by how cost and quality are prioritized in importance to the decision-maker. We emphasize that it is more important to select a solution that matches the prioritization of the objectives than to select the choice with the largest range of weights. However, in the vicinity of the desired weighting, selecting a robust choice can be advantageous if there is uncertainty in determining a single best choice of weights and if multiple decision-makers are involved in the process.

In addition to looking at which choices are best for different weights, examining the contributions to the desirability function from each criterion can help understand the trade-offs between different potential solutions. The trade-off plot in Fig. 7.4 shows the trade-off plot [Lu, Anderson-Cook, and Robinson (2011)] for the supplier example, with the inside axes showing the desirability score for each of cost and quality, while the outer axes show the original values. Again, note how the minimum (best) and maximum (worst) costs on the Pareto front have been assigned desirability scores of 1 and 0, respectively. Similarly the maximum and minimum quality values are mapped to 1 and 0. When combined with the plot in Fig. 7.3, we see that the suppliers who are best for high weightings of cost, like G and C, have excellent desirability scores for cost, but do poorly on quality. Suppliers who are selected as best for more balanced weighting of the two objectives have good desirability scores for both cost and quality.

The severity of the trade-off between cost and quality can be measured with the area under the Pareto front, after scaling with the desirability scores. Lu and Anderson-Cook (2013) describe the generalized version of this area, the hypervolume. We consider an area for 2 objectives (volume for 3, and hypervolume for 4+), which has maximum size of 1, since each dimension is scaled between 0 and 1. The idea is that if there are minimal trade-offs between objectives, then it is possible to find a solution quite close to the best possible for each objective individually, then we obtain an area close to 1. If improving one objective requires disproportionately large sacrifice from the other objective(s), then we obtain an area close to 0. Values close to 0.5 correspond to similar amounts of sacrifice to gain by moving between alternatives on the Pareto front. Figure 7.5 shows the scaled Pareto front



**Figure 7.4** Trade-off plot of suppliers best for at least one set of desirability function weights in Figure 7.3 with desirability scale shown on the inner axes and original cost and quality scores on outer axes.



**Figure 7.5** Area (hypervolume) under the Pareto front calculation for cost-quality example.

**TABLE 7.4 Central Composite Design with Three Responses**

| Natural Variables |         | Coded Variables |        | Responses     |                   |                          |
|-------------------|---------|-----------------|--------|---------------|-------------------|--------------------------|
| $\xi_1$           | $\xi_2$ | $x_1$           | $x_2$  | $y_1$ (yield) | $y_2$ (viscosity) | $y_3$ (molecular weight) |
| 80                | 170     | -1              | -1     | 76.5          | 62                | 2940                     |
| 80                | 180     | -1              | 1      | 77.0          | 60                | 3470                     |
| 90                | 170     | 1               | -1     | 78.0          | 66                | 3680                     |
| 90                | 180     | 1               | 1      | 79.5          | 59                | 3890                     |
| 85                | 175     | 0               | 0      | 79.9          | 72                | 3480                     |
| 85                | 175     | 0               | 0      | 80.3          | 69                | 3200                     |
| 85                | 175     | 0               | 0      | 80.0          | 68                | 3410                     |
| 85                | 175     | 0               | 0      | 79.7          | 70                | 3290                     |
| 85                | 175     | 0               | 0      | 79.8          | 71                | 3500                     |
| 92.07             | 175     | 1.414           | 0      | 78.4          | 68                | 3360                     |
| 77.93             | 175     | -1.414          | 0      | 75.6          | 71                | 3020                     |
| 85                | 182.07  | 0               | 1.414  | 78.5          | 58                | 3630                     |
| 85                | 167.93  | 0               | -1.414 | 77.0          | 57                | 3150                     |

for the cost–quality example, with the step-function between available solutions used to calculate the area. The step-function is recommended, since it is not possible to interpolate between adjacent solutions to find a “compromise,” and hence the area should reflect only those choices possible. The area under the Pareto front is 0.639, which shows that the trade-off between the objectives is not unduly severe.

**Example 7.2 A Three-Response Chemical Process** Often there are more than two objectives that are considered during the optimization, and they are the result of fitting response surface models based on a designed experiment. As an illustration, consider the experiment shown in Table 7.4. This is a central composite design in  $k = 2$  variables, conducted in a chemical process [see Montgomery (2013) for more details]. The two process variables are time and temperature, and there are three responses of interest:  $y_1 =$  yield,  $y_2 =$  viscosity, and  $y_3 =$  number-average molecular weight. The goal is to maximize yield, have viscosity near 65, and minimize molecular weight.

Simultaneous consideration of multiple responses involves first building an appropriate response surface model for each response and then trying to find a set of operating conditions that in some sense optimizes all responses or at least keeps them in desired ranges. We focus initially on the yield response  $y_1$ . Table 7.5 presents the output from Design-Expert for this response. From examining this table we notice that the software package first computes the sequential or extra sums of squares for the linear, quadratic, and cubic terms in the model (there is a warning message concerning aliasing in the cubic model because the CCD does not contain enough runs to support a full cubic model). Based on the small  $P$ -value for the quadratic term, we decided to fit the second-order model to the yield response. The computer output shows the final model in terms of both the coded variables

$$\hat{y}_1 = 79.94 + 0.99x_1 + 0.52x_2 - 1.38x_1^2 - 1.00x_2^2 + 0.25x_1x_2 \quad (7.3)$$

and the natural or actual factor levels

$$\hat{y}_1 = -1430.52 + 7.807\xi_1 + 13.271\xi_2 - 0.0551\xi_1^2 - 0.04005\xi_2^2 + 0.0100\xi_1\xi_2$$

**TABLE 7.5 Computer Output from Design-Expert for Fitting a Model to the Yield Response in Table 7.4**

| Response: Yield                        |                |    |              |                |                   |                  |
|--|----------------|----|--------------|----------------|-------------------|------------------|
| WARNING: The Cubic Model is Aliased!   |                |    |              |                |                   |                  |
| <b>Sequential Model Sum of Squares</b> |                |    |              |                |                   |                  |
| Source                                 | Sum of Squares | DF | Mean Square  | F-Value        | Prob > F          |                  |
| Mean                                   | 80062.16       | 1  | 80062.16     |                |                   |                  |
| Linear                                 | 10.04          | 2  | 5.02         | 2.69           | 0.1166            |                  |
| 2FI                                    | 0.25           | 1  | 0.25         | 0.12           | 0.7350            |                  |
| <i>Quadratic</i>                       | <i>17.95</i>   | 2  | 8.98         | <i>126.888</i> | <i>&lt;0.0001</i> | <i>Suggested</i> |
| Cubic                                  | 2.042E-003     | 2  | 1.021E-003   | 0.010          | 0.9897            | Aliased          |
| Residual                               | 0.49           | 5  | 0.099        |                |                   |                  |
| Total                                  | 80090.90       | 13 | 6160.84      |                |                   |                  |
| <b>Lack of Fit Tests</b>               |                |    |              |                |                   |                  |
| Source                                 | Sum of Squares | DF | Mean Square  | F-Value        | Prob > F          |                  |
| Linear                                 | 18.89          | 6  | 3.08         | 58.14          | 0.0008            |                  |
| 2FI                                    | 18.24          | 5  | 3.65         | 68.82          | 0.0006            |                  |
| <i>Quadratic</i>                       | <i>0.28</i>    | 3  | <i>0.094</i> | <i>1.78</i>    | <i>0.2897</i>     | <i>Suggested</i> |
| Cubic                                  | 0.28           | 1  | 0.28         | 5.31           | 0.0826            | Aliased          |
| Pure Error                             | 0.21           | 4  | 0.053        |                |                   |                  |



**TABLE 7.5 (Continued)**

| Model Summary Statistics   |                | R-Squared | Adjusted R-Squared | Predicted R-Squared | Press                              |
|--|----------------|-----------|--------------------|---------------------|------------------------------------|
| Source   | Std. Dev.      |           |                    |                     |                                    |
| Linear   | 1.37           | 0.3494    | 0.2193             | -0.0435             | 29.99                              |
| 2FI  | 1.43           | 0.3561    | 0.1441             | -0.2730             | 36.59                              |
| <i>Quadratic</i>   | 0.27           | 0.9828    | 0.9705             | 0.9184              | 2.35                               |
| Cubic  | 0.31           | 0.9828    | 0.9588             | 0.3622              | 18.33                              |
| Response: Yield  |                |           |                    |                     | <i>Suggested</i><br><i>Aliased</i> |
| ANOVA for Response Surface Quadratic Model Analysis of Variance Table [Partial Sum of Squares] |                |           |                    |                     |                                    |
| Source   | Sum of Squares | DF        | Mean Square        | F-Value             | Prob > F                           |
| Model  | 28.25          | 5         | 5.65               | 79.85               | <0.0001                            |
| A  | 7.92           | 1         | 7.92               | 111.93              | <0.0001                            |
| B  | 2.12           | 1         | 2.12               | 30.01               | 0.0009                             |
| A <sup>2</sup>   | 13.18          | 1         | 13.18              | 186.22              | <0.0001                            |
| B <sup>2</sup>   | 6.97           | 1         | 6.97               | 98.56               | <0.0001                            |
| AB   | 0.25           | 1         | 0.25               | 3.53                | 0.1022                             |
| <i>Residual</i>  | 0.50           | 7         | 0.071              |                     |                                    |
| <i>Lack of Fit</i>   | 0.28           | 3         | 0.094              | 1.78                | 0.2897                             |
| <i>Pure Error</i>  | 0.21           | 4         | 0.053              |                     |                                    |
| Cor Total  | 28.74          | 12        |                    |                     |                                    |
| Std. Dev.  | 0.27           |           |                    |                     |                                    |
| Mean   | 78.48          |           | R-Squared          | 0.9828              |                                    |
| C.V.   | 0.34           |           | Adj R-Squared      | 0.9705              |                                    |
| PRESS  | 2.35           |           | Pred R-Squared     | 0.9184              |                                    |
|  |                |           | Adeq Precision     | 23.018              |                                    |

(continued)

TABLE 7.5 (Continued)

| Factor         | Coefficient Estimate | DF | Standard Error | 95% CI Low | 95% CI High | VIF  |
|----------------|----------------------|----|----------------|------------|-------------|------|
| Intercept      | 79.94                | 1  | 0.12           | 79.66      | 80.22       |      |
| A-time         | 0.99                 | 1  | 0.094          | 0.77       | 1.22        | 1.00 |
| B-temp         | 0.52                 | 1  | 0.094          | 0.29       | 0.74        | 1.00 |
| A <sup>2</sup> | -1.38                | 1  | 0.10           | -1.61      | -1.14       | 1.02 |
| B <sup>2</sup> | -1.00                | 1  | 0.10           | -1.24      | -0.76       | 1.02 |
| AB             | 0.25                 | 1  | 0.13           | -0.064     | 0.56        | 1.00 |

Final Equation in Terms of Coded Factors:

$$\begin{aligned} \text{yield} = & +79.94 \\ & +0.99 * A \\ & +0.52 * B \\ & -1.38 * A^2 \\ & -1.00 * B^2 \\ & +0.25 * A * B \end{aligned}$$

Final Equation in Terms of Actual Factors:

$$\begin{aligned} \text{yield} = & -1430.52285 \\ & +7.80749 * \text{time} \\ & +13.27053 * \text{temp} \\ & -0.055050 * \text{time}^2 \\ & -0.040050 * \text{temp}^2 \\ & + 0.010000 * \text{time} * \text{temp} \end{aligned}$$

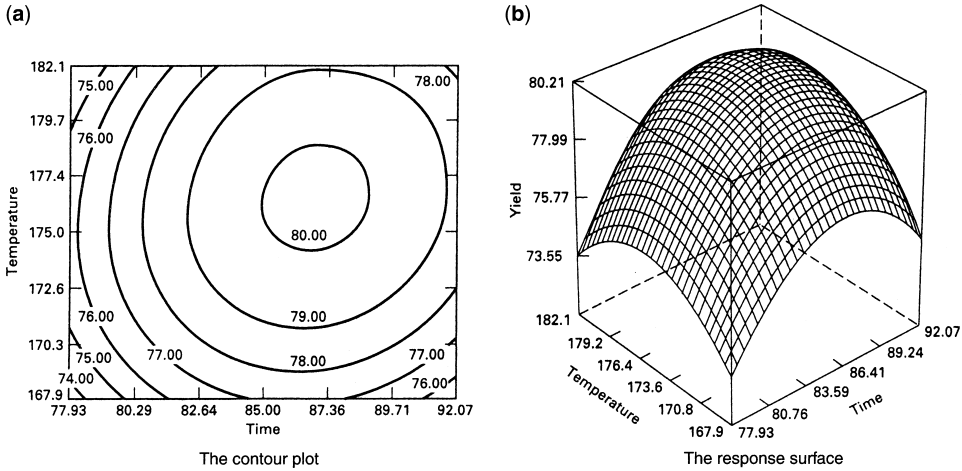
**TABLE 7.5 (Continued)**

| Diagnostics Case Statistics |                |              |                 |          |          |                  |                 |                  |  |  |
|-----------------------------|----------------|--------------|-----------------|----------|----------|------------------|-----------------|------------------|--|--|
| Run Order                   | Standard Order | Actual Value | Predicted Value | Residual | Leverage | Student Residual | Cook's Distance | Outlier <i>t</i> |  |  |
| 8                           | 1              | 76.50        | 76.30           | 0.20     | 0.625    | 1.213            | 0.409           | 1.264            |  |  |
| 6                           | 2              | 78.00        | 77.79           | 0.21     | 0.625    | 1.275            | 0.452           | 1.347            |  |  |
| 9                           | 3              | 77.00        | 76.83           | 0.17     | 0.625    | 1.027            | 0.293           | 1.032            |  |  |
| 11                          | 4              | 79.50        | 79.32           | 0.18     | 0.625    | 1.089            | 0.329           | 1.106            |  |  |
| 12                          | 5              | 75.60        | 75.78           | -0.18    | 0.625    | -1.107           | 0.341           | -1.129           |  |  |
| 10                          | 6              | 78.40        | 78.59           | -0.19    | 0.625    | -1.195           | 0.396           | -1.240           |  |  |
| 7                           | 7              | 77.00        | 77.21           | -0.21    | 0.625    | -1.283           | 0.457           | -1.358           |  |  |
| 1                           | 8              | 78.50        | 78.67           | -0.17    | 0.625    | -1.019           | 0.289           | -1.023           |  |  |
| 5                           | 9              | 79.90        | 79.94           | -0.040   | 0.200    | -0.168           | 0.001           | -0.156           |  |  |
| 3                           | 10             | 80.30        | 79.94           | 0.36     | 0.200    | 1.513            | 0.095           | 1.708            |  |  |
| 13                          | 11             | 80.00        | 79.94           | 0.060    | 0.200    | 0.252            | 0.003           | 0.235            |  |  |
| 2                           | 12             | 79.70        | 79.94           | -0.24    | 0.200    | -1.009           | 0.042           | -1.010           |  |  |
| 4                           | 13             | 79.80        | 79.94           | -0.14    | 0.200    | -0.588           | 0.014           | -0.559           |  |  |

\*\**Sequential Model Sum of Squares*\*\*: Select the highest-order polynomial where the additional terms are significant.

†*Lack of Fit Test*\*\*: Want the selected model to have insignificant lack-of-fit.

‡*Model Summary Statistics*\*\*: Focus on the model minimizing the "PRESS," or equivalently maximizing the "PRED R-SOR."



**Figure 7.6** Contour and response surface plots of the yield response in Table 7.4.

Figure 7.6 shows the three-dimensional response surface plot and the contour plot for the yield response in terms of the process variables time and temperature. It is relatively easy to see from examining these figures that the optimum is very near 176°F and 87 min of reaction time and that the response is at a maximum at this point. From examination of the contour plot, we note that the process may be slightly more sensitive to change in reaction time than to changes in temperature.

We obtain models for the viscosity and molecular weight responses ( $y_2$  and  $y_3$ , respectively) as follows:

$$\hat{y}_2 = 70.00 - 0.16x_1 - 0.95x_2 - 0.69x_1^2 - 6.69x_2^2 - 1.25x_1x_2 \tag{7.4}$$

$$\hat{y}_3 = 3386.2 + 205.1x_1 + 177.4x_2 \tag{7.5}$$

In terms of the natural levels of time ( $\xi_1$ ) and temperature ( $\xi_2$ ), these models are

$$\hat{y}_2 = -9028.78 + 13.394\xi_1 + 97.685\xi_2 - 0.0275\xi_1^2 - 0.2675\xi_2^2 - 0.05\xi_1\xi_2$$

and

$$\hat{y}_3 = -6308.0 + 41.021\xi_1 + 35.47\xi_2$$

Figures 7.7 and 7.8 present the contour and response surface plots for these models.

## 7.2 STRATEGIES FOR MULTIPLE RESPONSE OPTIMIZATION

We now outline a number of different approaches to selecting an optimal operating combination of the design factors. The different tools reflect that decision-makers do not all think about how to combine several objectives into a formal decision in the same way. The

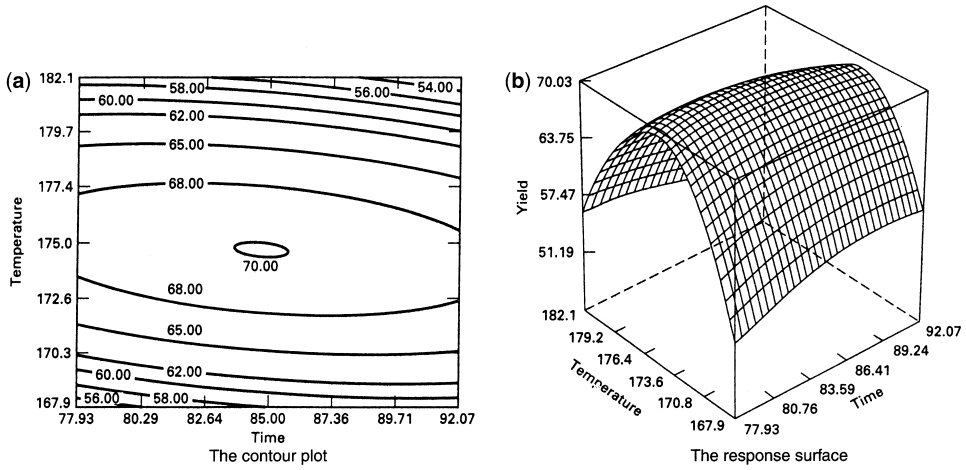


Figure 7.7 Contour plot and response surface plot of viscosity in Table 7.4.

tools also have varying ability to handle different numbers of design factors and responses efficiently.

### 7.2.1 Overlaying Contour Plots

A relatively straightforward approach to optimizing several responses that works well when there are only a few process variables is to **overlay the contour plots** for each response. [For an early illustration of this approach, see Lind, Goldin, and Hickman (1960).] Figure 7.9 shows an overlay plot for the three responses in Table 7.4, with contours for which  $y_1$  (yield)  $\geq 78.5$ ,  $62 \leq y_2$  (viscosity)  $\leq 68$ , and  $y_3$  (molecular weight  $M_n$ )  $\leq 3400$ . These are values that were judged to be close to the optimum for each response. If these boundaries represent important conditions that must be met by the process, then as the unshaded

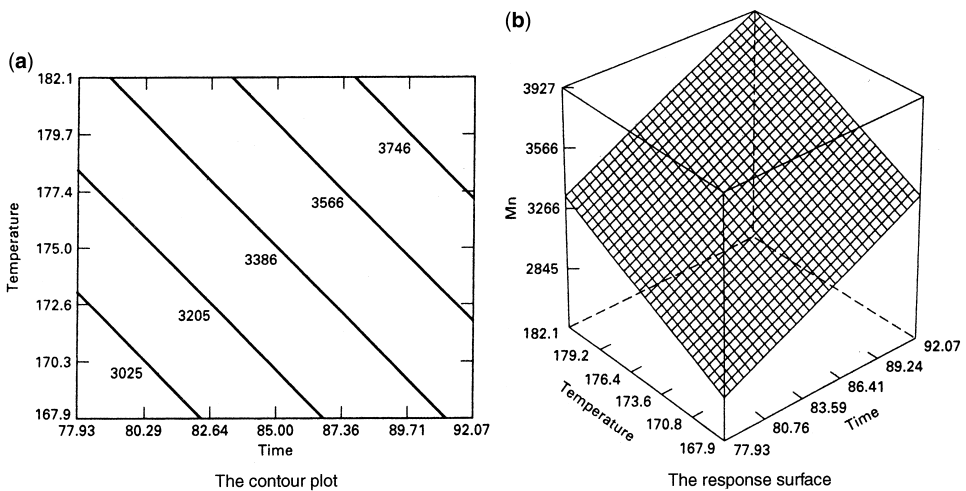
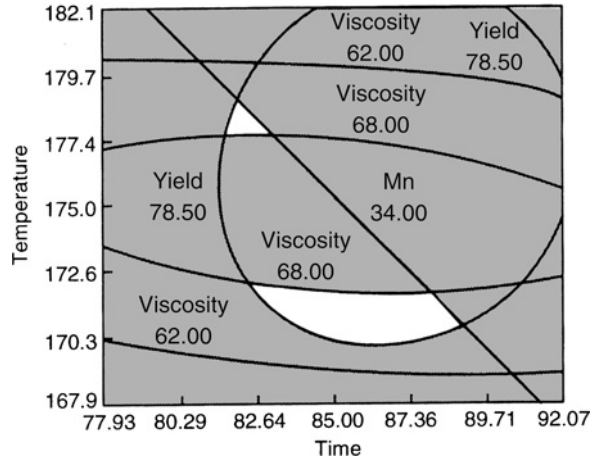


Figure 7.8 Contour and response surface plot of molecular weight in Table 7.4.



**Figure 7.9** Region of the optimum found by overlaying yield, viscosity, and molecular weight response surfaces in Table 7.4.

portion of Fig. 7.9 shows, there are a number of combinations of time and temperature that result in a satisfactory process. The experimenter can visually examine the contour plot to determine appropriate operating conditions, and select a region that is most desirable given other practical considerations. For example, it is likely that the experimenter would be most interested in the larger of the two feasible operating regions shown in Fig. 7.9.

When there are more than three design variables, overlaying contour plots becomes awkward, because the contour plot is two-dimensional, and  $k - 2$  of the design variables must be held constant to construct the graph. Often a lot of trial and error is required to determine which factors to hold constant and what levels to select to obtain the best view of the surface. Therefore, there is practical interest in more formal optimization methods for multiple responses.

**7.2.2 Constrained Optimization**

Another popular approach is to formulate and optimize as a constrained optimization problem. To illustrate, we might formulate the problem as

$$\begin{aligned} &\text{Max } y_1 \\ &\text{subject to } 62 \leq y_2 \leq 68 \\ &\quad y_3 \leq 3400 \end{aligned}$$

Here yield,  $y_1$ , was chosen as the primary focus with  $y_2$  and  $y_3$  chosen to be close enough to their ideal values. There are many numerical techniques that can be used to solve this problem. Sometimes these techniques are referred to as **nonlinear programming methods**. The Design-Expert software package solves this version of the problem using a direct search procedure. The two solutions found are

$$\text{Time} = 83.5 \quad \text{Temp} = 177.1 \quad \hat{y}_1 = 79.5$$

and

$$\text{Time} = 86.6 \quad \text{Temp} = 172.25 \quad \hat{y}_1 = 79.5$$

Notice that the first solution is in the upper (smaller) feasible region of the design space (refer to Fig. 7.9), whereas the second solution is in the larger region. Both solutions are very near to the boundary of the constraints.

In addition to direct search solutions, numerical optimization algorithms can be employed to find the optimum. Del Castillo and Montgomery (1993) describe this approach using the generalized reduced gradient (GRG) method. Carlyle et al. (2000) overview both direct search and numerical optimization techniques.

There are many ways to use nonlinear programming techniques to formulate and solve the multiple response optimization problem. Del Castillo (1996) describes an interesting approach involving constrained confidence regions. Optimal solutions are found that simultaneously satisfy confidence regions for the responses.

### 7.2.3 Desirability Functions

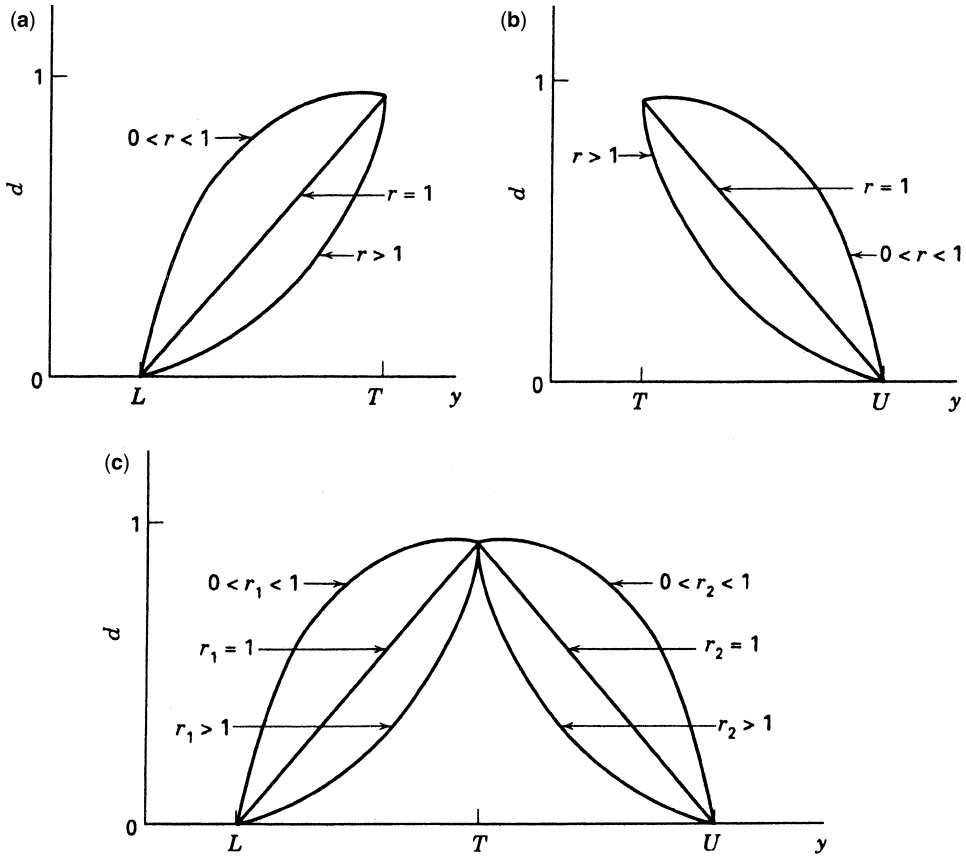
Introduced in Section 7.1, we return to the desirability function approach and provide some additional details about its implementation. To tailor the desirability to different prioritizations and to accommodate the ranges of options available for each response, several desirability function structures are available. The goal is to allow flexible specification of the desirability of alternatives and to adapt how quickly the desirability retreats from the optimal.

First, we consider the individual desirability functions for each response, shown in Fig. 7.10. If the objective of target  $T$  for the response  $y$  is a maximum value,

$$d = \begin{cases} 0, & y < L \\ \left(\frac{y-L}{T-L}\right)^r, & L \leq y \leq T \\ 1, & y > T \end{cases}$$

For the weight  $r = 1$ , the desirability function is linear, as shown in Fig. 7.10a. Choosing  $r > 1$  places more emphasis on being close to the target value, and choosing  $0 < r < 1$  makes this less important. If the target for the response is a minimum value (Fig. 7.10b), then we obtain

$$d = \begin{cases} 1, & y < T \\ \left(\frac{U-y}{U-T}\right)^r, & T \leq y \leq U \\ 0, & y > U \end{cases}$$



**Figure 7.10** Individual desirability functions for simultaneous optimization  $y$ . (a) Objective (target) is to maximize  $y$ . (b) Objective (target) is to minimize  $y$ . (c) Objective is for  $y$  to be as close as possible to the target.

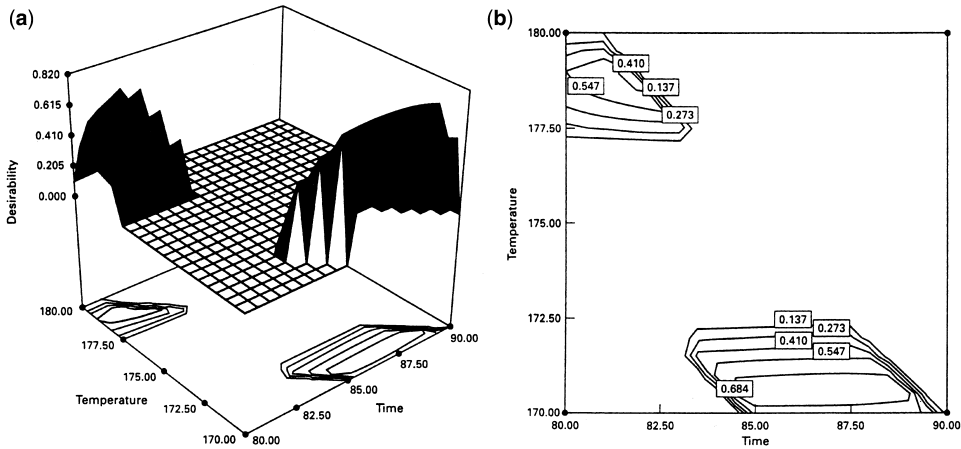
The two-sided desirability shown in Fig. 7.10c assumes that the target is located between the lower ( $L$ ) and upper ( $U$ ) limits, and is defined as

$$d = \begin{cases} 0, & y < L \\ \left(\frac{y-L}{T-L}\right)^{r_1}, & L \leq y \leq T \\ \left(\frac{U-y}{U-T}\right)^{r_2}, & T \leq y \leq U \\ 0 & y > U \end{cases}$$

Recall that the form of the desirability function to combine the different criteria can change how much good performance by one or more of the responses can overcome the penalty for poor performance of another. The additive form of the desirability in Equation 7.1 allows excellent performance of one response to overcome very poor performance of another response. The multiplicative form in Equation 7.2 protects more heavily from poor performance on any single response.

The Design-Expert software package was used to apply the desirability function approach to the chemical problem in Table 7.4. We chose  $T = 80$  as the target for the





**Figure 7.11** Desirability function response surface and contour plot for the problem in Table 7.4. (a) Response surface. (b) Contour plot.

yield response, with  $L = 70$ , and we set the weight for this individual desirability equal to unity. We set  $T = 65$  for the viscosity response with  $L = 62$  and  $U = 68$  (to be consistent with specifications), with both weights  $r_1 = r_2 = 1$ . Finally, we indicated that any molecular weight between 3200 and 3400 was acceptable. A multiplicative combining desirability function with equal weights for all three responses,  $D = (d_1 d_2 d_3)^{1/3}$ , was selected. Two solutions were found.

**Solution 1:**

|                    |                  |                    |
|--------------------|------------------|--------------------|
| Time = 86.1        | Temp = 170.3     | D = 0.929          |
| $\hat{y}_1 = 78.6$ | $\hat{y}_2 = 65$ | $\hat{y}_3 = 3319$ |

**Solution 2:**

|                    |                  |                    |
|--------------------|------------------|--------------------|
| Time = 80.3        | Temp = 179.2     | D = 0.855          |
| $\hat{y}_1 = 77.3$ | $\hat{y}_2 = 65$ | $\hat{y}_3 = 3400$ |

Solution 1 has the highest overall desirability. Notice that it results in on-target viscosity and acceptable molecular weight. This solution is in the larger of the two operating regions in Fig. 7.9, whereas the second solution is in the smaller region. Figure 7.11 shows a response surface and contour plot of the overall desirability function  $D$ .

Design-Expert uses direct search methods to maximize the desirability function  $D$ . Because the individual desirability functions are not differentiable, numerical optimization methods such as the GRG cannot be used directly. However, Del Castillo, Montgomery, and McCarville (1996) show how the individual desirability functions can be replaced with polynomial approximations that are everywhere differentiable so that the GRG and other derivative-based optimization algorithms can be employed.

### 7.2.4 Pareto Front Optimization

Introduced in Section 7.1, the Pareto front approach is commonly used in many engineering and operations research applications to eliminate noncontenders and to provide a smaller rational subset of candidates to consider in more detail.

The method consists of two sequential steps: (1) an objective step, where the Pareto front is constructed from available candidates and poor candidates (for which a solution on the Pareto front at least as good is available) are discarded as unworthy of further consideration, and (2) a subjective step, where solutions from the Pareto front are examined and compared to find the best match to the goals of the study. While the initial reaction to hearing about a step being subjective might sound undesirable, it is actually very advantageous to be able to tailor the choice of the best conditions under which to operate to match what is important to the decision-maker.

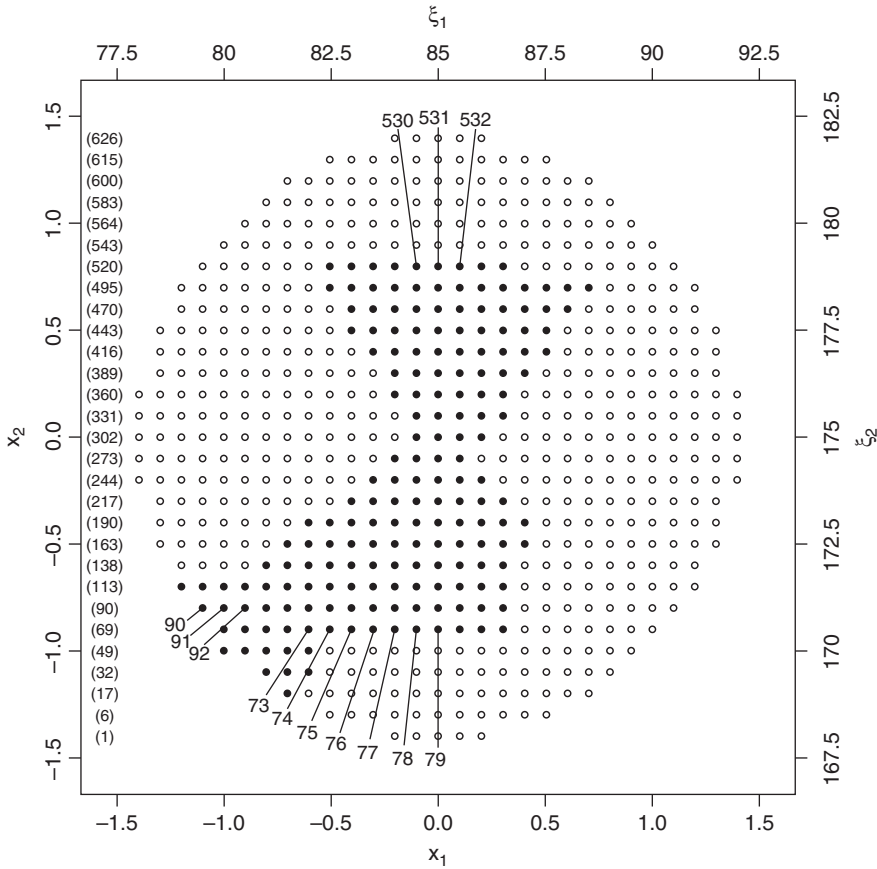
A solution is said to *Pareto dominate* another if all of its criteria values are at least as good, and at least one is strictly better. The *Pareto set* is the set of all solutions that are not Pareto-dominated by others, and the vectors of their response values form the Pareto front in the response space. The *Utopia point* is defined to be the set of response values which correspond the best available for each response. Hence, for the chemical process, these would correspond to the maximum estimated yield, minimum molecular weight and viscosity value closest to the target of 65. From looking at the overlaid contour plot in Fig. 7.8, it is clear that the best values for each response lie in different regions of the design space, and hence (as is most often the case) the Utopia point is not attainable. Geometrically, we can think of the Pareto set as the set of points on the outer edge of all obtainable solutions closest to the Utopia point.

Unlike the cost-quality optimization in Section 7.1 where the available choices were already enumerated, when we are optimizing over multiple estimated responses to obtain a best solution, we must first create a set of solutions with available response values from which we can identify the Pareto set. Summarizing results from Chapman et al. (2014a) for the two-factor, three-response chemical example, a grid of 630 points was created inside the circular design region defined by the central composite design. The range of values considered was inside a circle of radius  $\sqrt{2}$  for scaled time and temperature. The fineness of the grid was selected to be 0.1, to match what experimenters thought was the smallest discernible difference that could be implemented in production. Figure 7.12 shows the grid of points, with numeric labels (going left to right, row by row from the bottom) shown on the left side. The scaled time and temperature values are shown on the bottom and left axes, while the original scales are shown on the top and right.

The grid point coordinates are combined with Equations 7.3, 7.4, and 7.5 to obtain 630 sets of values that can be used to find the Pareto front, shown as the 181 solid circles (●) in Fig. 7.12. The remaining 449 locations shown with open circles (○) are not in the Pareto set and do not represent choices of design locations that should be considered further given the optimization objectives. Hence at the conclusion of the objective step of Pareto optimization, a substantial reduction in candidates to consider has already been achieved.

Since the decision-maker must ultimately decide on a single location on which to focus for future production, the subjective stage of Pareto optimization uses elements of the desirability function approach to highlight how different locations on the Pareto front perform for various prioritizations of the responses. Similar to Section 7.2.3, a multiplicative desirability functional form is selected, but here the scaling of the desirability of each response relative to the optimum was chosen based on the range of values from the Pareto front. The best observed value was scaled to 1, while the worst value on the Pareto front was scaled to 0.

A substantial advantage of the Pareto front approach is that with the reduced number of candidate locations under consideration, it is possible to examine all possible weight



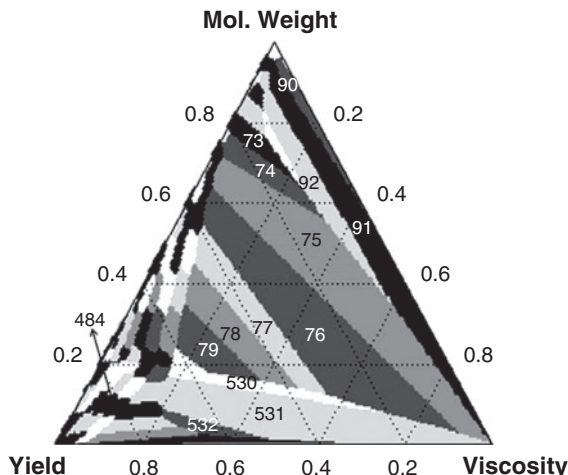
**Figure 7.12** Grid points in design space used to approximate yield, viscosity, and molecular weight values. Solid circles are design locations on the Pareto front.

combinations of how to value the different responses. Figure 7.13 shows a mixture plot of all possible weight combinations for the desirability function

$$D = d_{yield}^{w_{yield}} d_{Visc}^{w_{Visc}} d_{MW}^{w_{MW}} \tag{7.6}$$

where  $w_{yield} + w_{Visc} + w_{MW} = 1$ . Details about mixtures are given in Chapter 12, but the key features of the mixture plot for comparing the performance of different weightings of the response optimizations are as follows: Vertices denote assigning 100% of the weight to a single response. For example, the top vertex shows the best location when all weight is assigned to molecular weight. Edges assign nonzero weight to the two responses on that edge, with a zero weight assigned to the other response. Interior points have nonzero weights for all three responses. Moving closer to a vertex increases the weighting of that response.

From Fig. 7.13, we see that for near-equal weighting of the three responses (near the centroid of the triangle), locations 75–79 provide the best balance of trade-offs between responses. These locations correspond to scaled temperatures of  $-0.9$  and scaled times

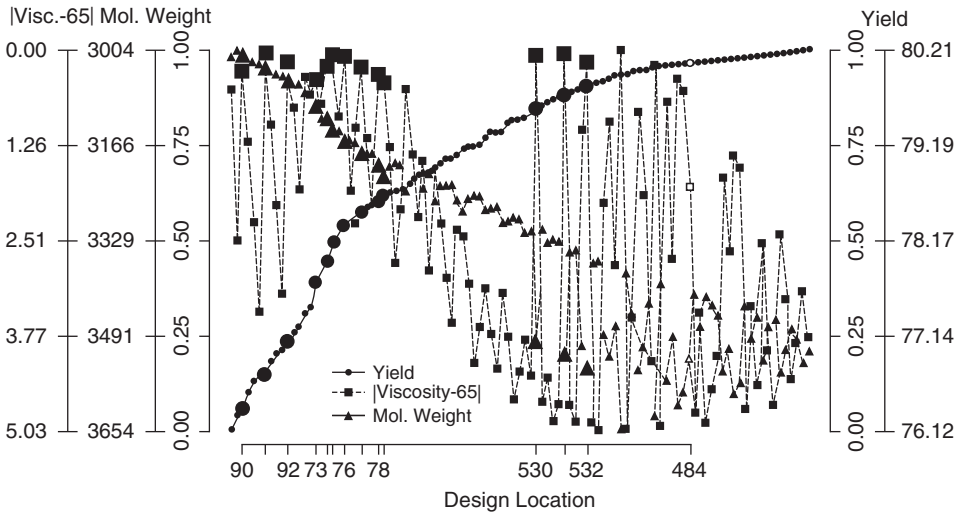


**Figure 7.13** Mixture plot identifying best design locations in Fig. 7.12 for all possible weight combinations of response desirability scores.

ranging from  $-0.4$  to  $0$ . Table 7.6 shows some of the design space locations that are best for larger proportions of the weighting space. By examining the set of all possible weights, it allows the decision-maker to have a better understanding of which regions perform well for different priorities. Figure 7.14 shows the trade-off plot of different solutions sorted by yield, with the inner axes showing the desirability score and the outer axes showing the original response values. From this plot it is clear why the locations between 75 and 79 are best for nearly equal weighting of the responses. Most other solutions have at least

**TABLE 7.6 Promising Design Locations with Their Estimated Responses and Percentage of Mixture Area**

| Location | Coded Variables |       | Natural Variables |         | Estimated Responses |             |             | % Area |
|----------|-----------------|-------|-------------------|---------|---------------------|-------------|-------------|--------|
|          | $x_1$           | $x_2$ | $\xi_1$           | $\xi_2$ | $\hat{y}_1$         | $\hat{y}_2$ | $\hat{y}_3$ |        |
| 73       | -0.6            | -0.9  | 82                | 170.5   | 77.70               | 64.61       | 3103.48     | 1.44%  |
| 74       | -0.5            | -0.9  | 82.5              | 170.5   | 77.93               | 64.78       | 3123.99     | 2.38%  |
| 75       | -0.4            | -0.9  | 83                | 170.5   | 78.14               | 64.94       | 3144.47     | 16.54% |
| 76       | -0.3            | -0.9  | 83.5              | 170.5   | 78.31               | 65.08       | 3164.99     | 16.68% |
| 77       | -0.2            | -0.9  | 84                | 170.5   | 78.46               | 65.21       | 3185.50     | 5.07%  |
| 78       | -0.1            | -0.9  | 84.5              | 170.5   | 78.57               | 65.33       | 3206.03     | 4.43%  |
| 79       | 0               | -0.9  | 85                | 170.5   | 78.66               | 65.44       | 3226.54     | 3.04%  |
| 90       | -1.1            | -0.8  | 79.5              | 171     | 76.35               | 64.72       | 3018.62     | 1.26%  |
| 91       | -1              | -0.8  | 80                | 171     | 76.72               | 64.95       | 3039.13     | 8.16%  |
| 92       | -0.9            | -0.8  | 80.5              | 171     | 77.05               | 65.16       | 3059.69     | 3.70%  |
| 484      | 0.2             | 0.6   | 86                | 178     | 80.06               | 66.81       | 3533.60     | 0.29%  |
| 530      | -0.1            | 0.8   | 84.5              | 179     | 79.58               | 65.07       | 3507.61     | 1.87%  |
| 531      | 0               | 0.8   | 85                | 179     | 79.71               | 64.96       | 3528.05     | 11.82% |
| 532      | 0.1             | 0.8   | 85.5              | 179     | 79.82               | 64.84       | 3548.63     | 1.10%  |

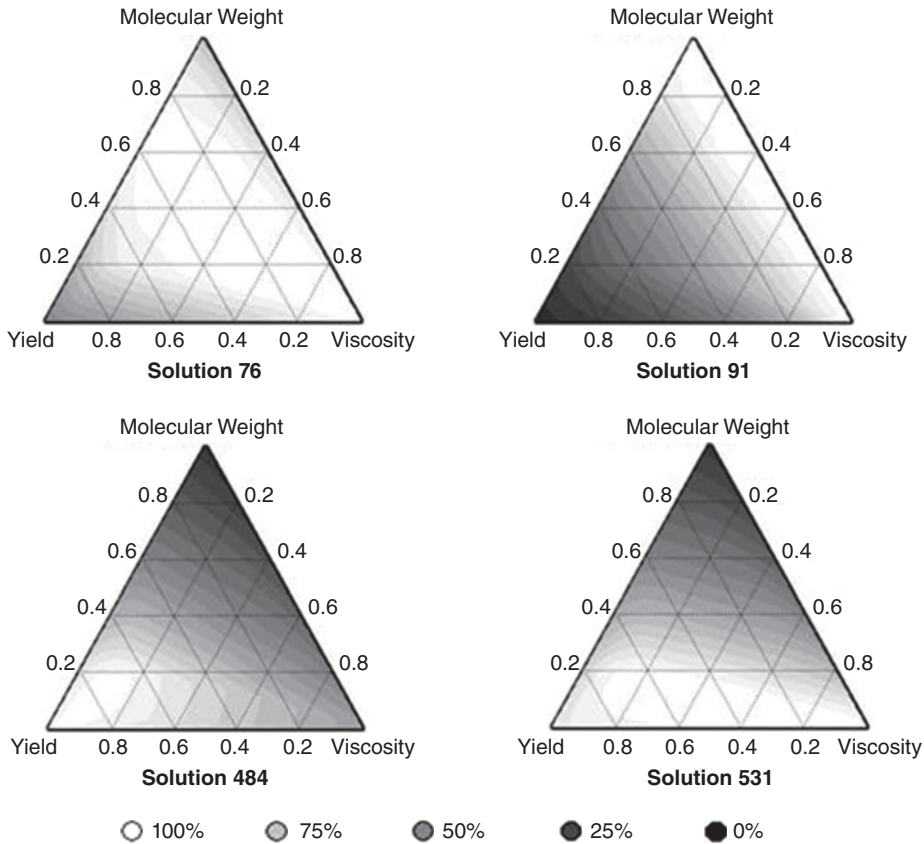


**Figure 7.14** Trade-off plot of design locations that are best for at least one weight combination of the desirability scores of yield, viscosity and molecular weight.

one desirability score below 0.5, while for the 75–79 locations there are no very poor performers. For other locations in the plot, there are one or two good values, but coupled with at least one poor value.

Recall that the overall goal is to identify a single location in the design space at which future production will run. Hence, once we have identified several candidate locations, it is helpful to understand their performance better. The synthesized efficiency plot introduced in Lu and Anderson-Cook (2012) looks at how individual locations perform relative to the best available at any given weight combination. In Section 7.1 we considered the synthesized efficiency value for a single weight combination, while now we expand this to look at all of these values across the various weight combinations. Figure 7.15 shows the synthesized efficiency plots for 4 locations that are best for very different weight combinations in Figure 7.13. For a fixed weight in Equation 7.6, the desirability function value for the shown location is divided by the best available desirability function value from any location, and this calculation is replicated for all possible weight combinations. An efficiency of 100% means that this design location equals the best possible. Shades of white, gray, and black show deterioration of performance relative to the best possible, with each shade denoting a drop of 5%. Hence solid white means that the chosen location is at least 95% efficient. The white region for location 76 indicates that is very close to best for large proportions of the weightings. Location 91 is best when molecular weight and/or viscosity are weighted heavily. Location 484 is best for heavy weighting of yield, while location 531 favors strong weighting of yield or viscosity. The dark regions indicate that these locations perform poorly for some weight combinations and should be avoided if these weights represent important prioritizations of the responses.

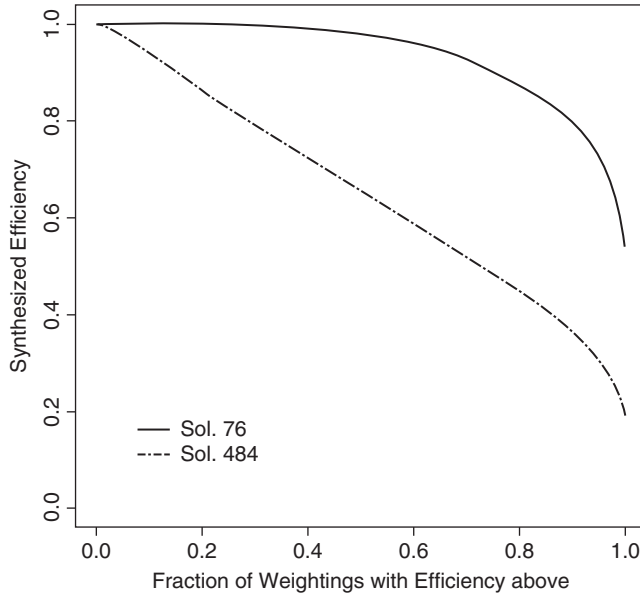
If near-equal weighting of the responses is desired, then location 76 with coordinates  $(x_1, x_2) = (-0.3, -0.9)$  might be a good choice. If yield were deemed more important, then 484 with  $(x_1, x_2) = (0.2, 0.6)$  or some location near it would provide good performance. To summarize the performance of the selected location across the range of different weight



**Figure 7.15** Synthesized efficiency plots for some selected design locations.

combinations, the fraction of weight (FWS) space plot [Lu et al. (2013)] can be used. The plot shown in Fig. 7.16 provides an overall, quantitative summary of the performance of an individual location across the entire weighting space and can be used to easily compare competing solutions. The line for each location plots the fraction of the weighting space in the triangle shown in Fig. 7.13 for which the location’s synthesized efficiency is at least the specified percentage. To calculate the values for the curve, the desirability function values for all sets of weights are divided by the best available desirability function value at that weight combination.

From Fig. 7.16, it is clear that location 76 has the more robust performance across the entire range of weight combinations, but this does not mean that this is the better location to select. The highest priority should be to match the objectives of the optimization to the weight combination, and select based on this. However, the FWS plot can be helpful for showcasing the robustness of the alternatives across a range of weight combinations of interest. The FWS plot can also be used to summarize the synthesized efficiency for just a portion of the weight space [Lu, Chapman, Anderson-Cook (2013)]. An alternate version of the plot could show the efficiencies relative to the best available locations for just the portion of the weight space of interest. For example, the FWS plot for location 76 could highlight performance for only those weight combinations near equal weighting of the three responses.



**Figure 7.16** Fraction of weight space plot for locations 76 and 484 in Fig. 7.12.

Hence, the Pareto front approach provides a two-step process for eliminating inferior choices from further consideration, and then a variety of graphical tools to allow exploration and understanding of performance for different priorities. If several stakeholders are involved in the decision-making process, having tools to visualize the trade-offs between alternatives can be helpful to facilitate compromise and consensus. The Pareto front approach can be applied for any number of responses, although the number of solutions found tends to increase quickly if many responses are considered. In addition, the visualization tools become more difficult to adapt with more than four responses.

### 7.2.5 Other Options for Optimization

In addition to the approaches illustrated above, several creative ideas have been put forth in the literature, and some enjoy considerable use in industry. We give a brief overview of these other methods.

**Dual Response Approach** In the case of two responses, a useful technique termed the **dual response approach** was introduced by Myers and Carter (1973). It is assumed that the two responses can be categorized as **primary** and **secondary**. The goal then involves the determination of conditions  $\mathbf{x}$  on the design variables that produces  $\max \hat{y}_p(\mathbf{x})$  [or  $\min \hat{y}(\mathbf{x})$ ] subject to a constraint on  $\hat{y}_s(\mathbf{x})$ . Here  $\hat{y}_p(\mathbf{x})$  is the primary response function and  $\hat{y}_s(\mathbf{x})$  is the secondary response function. The methodology, which is not unlike ridge analysis discussed in Chapter 6, produces a locus of coordinates in which various values of  $\hat{y}_s(\mathbf{x})$  are considered. For example, in a yield–cost scenario one might let the secondary response be yield and determine conditions on  $\mathbf{x}$  that minimize cost for fixed satisfactory levels of yield. This approach was the forerunner of more sophisticated nonlinear programming procedures. Biles (1975) and Del Castillo and Montgomery (1993) discuss the extension

of this concept to more than two responses. Del Castillo, Fan, and Semple (1997) discuss computing global optima in dual-response systems, and Fan (2000) gives a Fortran program. Also see Lin and Tu (1995).

**Khuri–Conlon Approach** Khuri and Conlon (1981) introduced a procedure that is based on a **distance** function that computes the overall closeness of the response functions to their respective optima at the same set of conditions. That is, it gives a measure of closeness to the *ideal optimum*. One then finds conditions on  $\mathbf{x}$  that minimize this distance function over all  $\mathbf{x}$  in the experimental region. The distance function they use is

$$[\hat{\mathbf{y}}(\mathbf{x}) - \boldsymbol{\theta}]' \sum_{\hat{\mathbf{y}}(\mathbf{x})}^{-1} [\hat{\mathbf{y}}(\mathbf{x}) - \boldsymbol{\theta}]$$

where  $\hat{\mathbf{y}}(\mathbf{x})$  is the vector of estimated responses at the point  $\mathbf{x}$ ,  $\Sigma_{\hat{\mathbf{y}}(\mathbf{x})}$  is the covariance matrix for the estimated responses at this location, and  $\boldsymbol{\theta}$  is the vector of target values or ideal optimal values for the responses. The optimum value of  $\mathbf{x}$  would minimize this distance function. A very nice aspect of this approach is that it considers directly the correlation structure of the responses and the quality of the response predictions. However, it does not allow the analyst to create appropriate priorities for the individual responses.

**Squared Error Loss Approaches** Several authors have recommended approaches based on squared error loss (indeed, the Khuri–Conlon method is a squared error loss procedure). For example, see Pignatiello (1993), Ames et al. (1997), and Vining (1998). In Vining's approach the squared error loss function is

$$L = [\hat{\mathbf{y}}(\mathbf{x}) - \boldsymbol{\theta}]' \mathbf{C} [\hat{\mathbf{y}}(\mathbf{x}) - \boldsymbol{\theta}]$$

where  $\mathbf{C}$  is a positive definite matrix of weights or costs. He minimizes the expected loss

$$E(L) = \{E[\hat{\mathbf{y}}(\mathbf{x})] - \boldsymbol{\theta}\}' \mathbf{C} \{E[\hat{\mathbf{y}}(\mathbf{x})] - \boldsymbol{\theta}\} + \text{Trace} \left[ \mathbf{C} \sum_{\hat{\mathbf{y}}(\mathbf{x})} \right]$$

The estimated expected loss is

$$\widehat{E(L)} = [\hat{\mathbf{y}}(\mathbf{x}) - \boldsymbol{\theta}]' \mathbf{C} [\hat{\mathbf{y}}(\mathbf{x}) - \boldsymbol{\theta}] + \text{Trace} \left[ \mathbf{C} \sum_{\hat{\mathbf{y}}(\mathbf{x})} \right]$$

There are several different choices for the matrix  $\mathbf{C}$ .

### 7.3 A SEQUENTIAL PROCESS FOR OPTIMIZATION—DMRCS

To this point, we have described different tools for the optimization process. We now consider how to organize a progression of steps into making a structured decision that is defensible and rational. A five-step process was described in Anderson-Cook and Lu (2015) to evaluate and balance competing objectives in a structured and defensible way, while still allowing the flexibility to incorporate and evaluate different subjective weightings of the facets of the decision. Similar to the DMAIC (Define–Measure–Analyze–Improve–Control) process popularized by Six Sigma, this complex process is broken into distinct



steps and highlight some of the important aspects of each. The five steps are: Define–Measure–Reduce–Combine–Select.

**DEFINE**—Like with the DMAIC process, the first step is to clarify the scope of the optimization by identifying the problem to be solved, the appropriate responses to consider, and what the appropriate design factors and ranges over which to optimize. This step is critical for framing the key components of the decision to ensure that the correct problem is solved. For optimizing multiple responses, it is important to make sure that all key facets of the decision are represented with an appropriate response. It is important to think broadly about diverse responses that are important to the decision, and not have multiple metrics for similar dimensions.

**MEASURE**—The second step that is also common to the DMAIC process is measure. This focuses on the need to be confident that we have high-quality data on which to base the decision. The facets of the decision, here the responses from the designed experiment, need to be precisely and accurately measured to allow fair and consistent comparisons between potential solutions. In addition, it is beneficial to verify that the setting of the input factor levels is done as precisely as possible. If high-quality quantitative measures of the responses cannot be used, then the quality of the decision may be compromised.

**REDUCE**—The third step focuses on reducing the optimization in several different dimensions. First, optimizing based on too many responses may lead to mediocre values for any response individually, as the trade-offs between responses generally increase as the number of facets considered increases. Hence, some triage of the response may lead to reduction in the number considered, as highly correlated responses or those measuring secondary objectives may be set aside. The second type of reduction to consider is to eliminate noncontending solutions from further consideration. The Pareto front described in Section 7.2.4 is an objective and rational way to include only those options that are suitable for further consideration. Trimming the number of solutions considered allows more meaningful exploration of the alternatives, but it is reassuring to use an approach that will not spuriously eliminate any worthwhile choices.

**COMBINE**—Next, we consider ways of combining the different metrics, often measured on different scales, into comparable forms. Many of tools described in Section 7.2 represent alternatives that merge the Reduce and Combine steps into a single operation. Overlaying contours, constrained optimization, or the desirability function approach with specific weights all provide ways of eliminating portions of the design space, and narrowing the focus to more manageable numbers of alternatives, based on a specified description of how to combine the multiple responses into a self-contained optimization goal.

The Pareto front approach separates these two steps into objective and subjective steps, with the construction of the front objectively eliminating noncontenders, while the subjective step allows for prioritizing these choices based on the goals of the optimization and the importance of the different responses. The Combine step allows examination of leading contenders while incorporating and evaluating our priorities for a good final solution. Using appropriately scaled individual desirability functions and a functional form for combining these into a single expression is the key to understanding the important trade-offs. An advantage of the Pareto front approach is that it has straightforward tools for examining the robustness of the result across several different priorities with quick examination of the impact of different weighting, scaling, and metric forms on the decision.

**SELECT**—At the conclusion of the Select step, the goal is to have identified a best location in the design space that suitably optimizes the response values in a way that best

matches the priorities of the study. Using graphical tools to help understand the relative merits and weaknesses of the alternatives can help build a team consensus if the goals of multiple decision-makers are divergent. In addition, it provides tools to formally justify the particular choice. If several choices perform similarly, an effective way of choosing between these options is to compare their performance on the supplementary criteria, such as cost and ease of implementation.

Once a final decision has been made, it is appropriate to verify that the results obtained from a particular combination of design factor levels yield the desired results in full-scale operation.

The use of a structured approach to optimization can greatly help with both its implementation and the quality of results obtained. Omitting careful definition of the optimization to match study goals, imprecise measurement of the inputs or response, and too much of a “black box” approach to optimization can lead to results that are not a good match for the overall intent of the process.

#### 7.4 INCORPORATING UNCERTAINTY OF RESPONSE PREDICTIONS INTO OPTIMIZATION

In Section 7.2.4 solutions using the Pareto front approach were identified for the two factor chemical process optimization for yield, viscosity and molecular weight. However, the optimization was based on the point estimates for the responses from Equations 7.3, 7.4, and 7.5. These equations do not take into account the natural variability associated with each response, and hence the predicted values for each of the responses as the selected optimum design location might not be attained in production when implemented.

We now consider two alternatives for optimization which take into account the natural variability in the responses, and how they might impact our choice of where to operate. The first approach is a conservative one that replaces the point estimates of the responses with the worst-case result based on the response’s prediction interval at any location. Recall from Section 2.6 that the equation for a prediction interval of a new observation using the model parameter estimates is

$$\hat{y}(x_0) \pm t_{\frac{\alpha}{2}, n-p} \sqrt{\hat{\sigma}^2 (1 + x_0'(\mathbf{X}'\mathbf{X})^{-1}x_0)}$$

where  $\hat{y}(x_0) = x_0'\hat{\beta}$  is the point estimate from Equation 7.3–7.5 for each of the responses at a given design location, and  $t_{\frac{\alpha}{2}, n-p}$  is a quantile from the  $t$ -distribution. Suppose interest lies in basing results on the 95% prediction intervals for the responses. First, we define the worst-case outcome for each of the three responses. Since we want to maximize yield, the least desirable outcome is the lower bound of the 95% prediction interval, since this is furthest away from the maximum. Similarly, for minimizing molecular weight, we select the upper bound of the prediction interval to be conservative. For viscosity, where the goal is to be as close to the target of 65 as possible, we find the maximum distance from 65, by looking at how far either the upper or lower bound is from the target.

These values are obtained for each response for all 630 grid locations shown in Fig. 7.12, and they are used as the basis for constructing the Pareto front. Details of this optimization are given in Chapman et al. (2014a), with the grid locations identified as in the Pareto

set being quite similar to those selected with the point estimates. However, depending on the relative size of the natural variability between the different responses, the Pareto front locations could change. With this optimization, the values identified for the chosen location in the design space correspond to the worst-case bound of the 95% prediction interval values. Hence when the solution is implemented, the user should be confident in at least attaining these values, with a good probability of exceeding them.

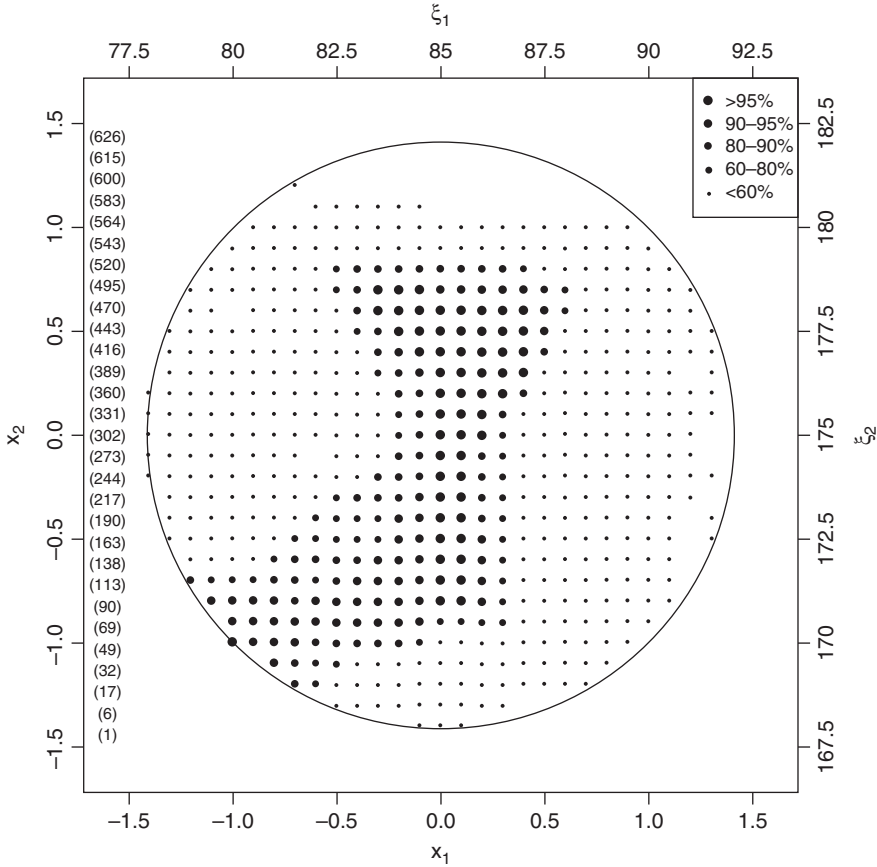
A second, more comprehensive approach is described in Chapman et al. (2014b), which uses simulation of the estimated model parameters for each of the responses to characterize performance of the Pareto front and its optimizing results. In the initial stage, the results of the regression analysis shown in Table 7.5 and corresponding versions for the other responses are used to generate a large number of sets of model parameter values consistent with the observed data. These values are then used to obtain response surfaces representative of the plausible relationship between the inputs and the responses. Recall from Chapter 2 that for each of the responses the model parameter estimates have a Normal distribution centered at the true but unknown model parameters,  $\hat{\beta} \sim MVN(\beta, \sigma^2(X'X)^{-1})$  and that we can estimate the variability with  $\hat{\sigma}^2 = MSE = \frac{SSE}{n-p}$ .

To generate new response surfaces, we repeatedly simulate new  $\hat{\beta}^*$ 's consistent with the model parameter estimates, with the following process:

1. Identify and fit the appropriate model of the form  $y_r = X_r\beta_r + \epsilon_r$  for each response variables.
2. For each of the  $m$  response variables, randomly generate  $\hat{\beta}_r^*$  from  $MVN(\hat{\beta}_r, \hat{\sigma}_r^2(X_r'X_r)^{-1})$ ,  $r = 1, \dots, m$ .
3. For each of the  $m$  response variables, use the simulated coefficients to approximate the response surface over a grid of points in the design region of interest. In the chemical process example, we use a grid of 630 locations inside the circle of radius  $\sqrt{2}$ . This collection of response surfaces serves as the basis for understanding of the impact of estimation uncertainty.

In the next objective stage, the goal is to summarize how the uncertainty from individual responses propagates to changes in the Pareto fronts. Using the fronts for each simulated surface, a variation of the plot shown in Fig. 7.12 can be used to summarize the frequency with which design locations appear on the front. Combinations that appear infrequently can be eliminated from further consideration. Figure 7.17 shows how often different points in the grid appear on the collection of 500 simulated Pareto fronts, with larger points corresponding to most frequently chosen locations. Note that the general shape of the locations identified matches those identified when using the  $\hat{\beta}$  point estimates, but with some calibration of how often we might expect those locations to be selected if the same experiment were repeated and slightly different results obtained.

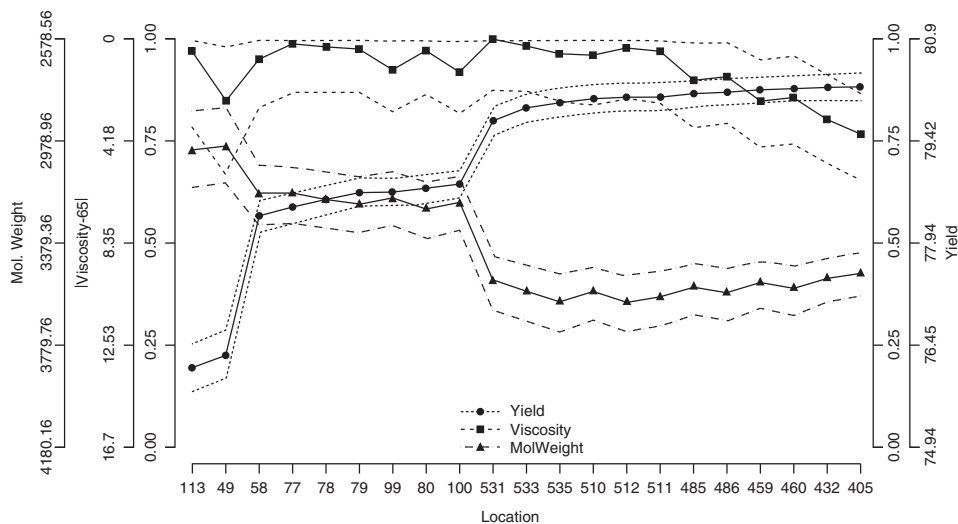
With an understanding of the Pareto front, the next step involves incorporating subjective choices about the relative weighting of the different objectives and making a final selection of a best location to operate to match the study goals. This process can be broken into several substeps, including: (a) highlighting promising general solutions that are robust across a larger number of weight combinations, (b) finding promising solutions for a more focused region of the weight space, and (c) making a final choice after comparison of competitive alternatives.



**Figure 7.17** Grid points in design space used to approximate yield, viscosity and molecular weight values, with the size of the circle proportional to the frequency with which the location appears on Pareto fronts in the simulation.

To start this step, a desirability function for converting each response to the desirability score is selected, where now it is recommended to consider the complete range of observed values from throughout the design space, as well as across all the simulated sets of responses. With each of these individual desirability functions, a desirability function form for combining them into a single measure is selected. Again, we select the multiplicative form for the chemical process example. An adapted trade-off plot is shown in Fig. 7.18 for some of the solutions with larger areas in Fig. 7.13 with uncertainty intervals obtained from the simulated values. This helps highlight what results might be expected under regular operating conditions.

Once more specific weight combinations are selected to match the overall goal of the study, then a more detailed examination of the alternatives can be performed. If weight combinations with near-equal weights for the desirability of all three responses are of interest, then Table 7.7 shows the most frequently selected location across all the simulation for a selected weight combinations with weights between 0.2 and 0.4 for all responses. The locations for a given label number are shown in Fig. 7.17. The choice for equal weights for all responses is shown in bold.



**Figure 7.18** Trade-off plot of three responses with associated uncertainty derived from simulated responses.

Alternatively, if weighting the desirability of yield is more important, then results for the most commonly selected location in the design space for these priorities can be found. Table 7.8 shows the most commonly selected locations when yield is weighted between 60 and 80% and the other responses are weighted between 10 and 20%. Note how which region of the design space is preferred changes with different prioritizations. Chapman et al. (2014b) provide additional details about how these choices can be compared and evaluated.

To gain understanding about how the selected location performs for different weight combinations, we can adapt the FWS plot shown in Fig. 7.16 for the analysis without

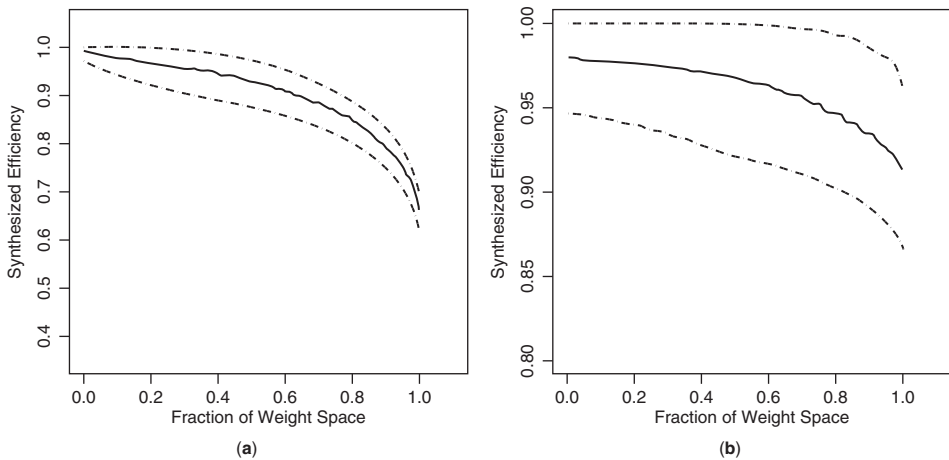
**TABLE 7.7 Most Commonly Selected Design Location for Different Weight Combinations with Near-Equal Weighting of Response Desirability**

| Weight Vector      | Top Location | Weight Vector             | Top Location |
|--------------------|--------------|---------------------------|--------------|
| (0.33, 0.4, 0.27)  | 79           | (0.23, 0.4, 0.37)         | 76           |
| (0.37, 0.37, 0.27) | 80           | (0.27, 0.33, 0.4)         | 77           |
| (0.37, 0.4, 0.23)  | 508          | (0.27, 0.37, 0.37)        | 77           |
| (0.4, 0.33, 0.27)  | 508          | (0.27, 0.4, 0.33)         | 77           |
| (0.4, 0.37, 0.23)  | 509          | (0.3, 0.3, 0.4)           | 77           |
| (0.4, 0.4, 0.2)    | 509          | (0.3, 0.33, 0.37)         | 77           |
| (0.33, 0.27, 0.4)  | 78           | (0.3, 0.37, 0.33)         | 78           |
| (0.37, 0.23, 0.4)  | 100          | (0.3, 0.4, 0.3)           | 78, 79       |
| (0.37, 0.27, 0.37) | 78           | (0.33, 0.3, 0.37)         | 78           |
| (0.4, 0.2, 0.4)    | 100          | <b>(0.33, 0.33, 0.33)</b> | <b>78</b>    |
| (0.4, 0.23, 0.37)  | 101          | (0.33, 0.37, 0.3)         | 79           |
| (0.4, 0.27, 0.33)  | 101          | (0.37, 0.3, 0.33)         | 79           |
| (0.2, 0.4, 0.4)    | 76           | (0.37, 0.33, 0.3)         | 78           |
| (0.23, 0.37, 0.4)  | 76           | (0.4, 0.3, 0.3)           | 101          |

**TABLE 7.8 Most Commonly Selected Design Location for Different Weight Combinations with Higher Weighting of Yield Desirability**

| Weight Vector      | Top Location |
|--------------------|--------------|
| (0.73, 0.1, 0.17)  | 405          |
| (0.73, 0.13, 0.13) | 432          |
| (0.73, 0.17, 0.1)  | 460, 486     |
| (0.77, 0.1, 0.13)  | 405          |
| (0.77, 0.13, 0.1)  | 432          |
| (0.8, 0.1, 0.1)    | 405          |
| (0.7, 0.1, 0.2)    | 376          |
| (0.7, 0.13, 0.17)  | 432          |
| (0.7, 0.17, 0.13)  | 432          |
| (0.7, 0.17, 0.13)  | 432          |
| (0.6, 0.2, 0.2)    | 485          |
| (0.63, 0.17, 0.2)  | 432          |
| (0.63, 0.2, 0.17)  | 485          |
| (0.67, 0.13, 0.2)  | 376          |
| (0.67, 0.17, 0.17) | 432          |
| (0.67, 0.2, 0.13)  | 486          |

taking into account the uncertainty. Figure 7.19 shows the FWS plot for location 77 with coordinates  $(x_1, x_2) = (-0.2, -0.9)$  (one that shows up frequently in Table 7.7) for the entire range of weight combinations, as well as for just those weight combinations where the weights for all response desirabilities are between 0.2 and 0.4. Note that there are now uncertainty values associated with the curves, which summarize the range of synthesized efficiencies observed across the simulations.



**Figure 7.19** Fraction of weight space plot for location 77 with coordinates  $(x_1, x_2) = (-0.2, -0.9)$  for (a) the entire range of weight combinations and (b) weight combinations where weights for all responses are between 0.2 and 0.4.

In this example, the highlighted design locations are relatively similar between the original analysis shown in Section 7.2.4 and those shown in this section. For the weight combinations with each weight constrained between 0.2 and 0.4 for all three responses, adjacent locations are identified. When we look at the predicted responses directly, the best location might be chosen to be location 76 with coordinates  $(x_1, x_2) = (-0.3, -0.9)$ , while when we take into account uncertainty, location 77 with coordinates  $(x_1, x_2) = (-0.2, -0.9)$  seems to perform slightly better.

The similarity in best locations for a particular goal is in part due to the relative narrow differences in the variability between the various responses. To understand where incorporating the uncertainty in the responses might have the greatest impact, Chapman et al. (2015) summarize several multiple interdependent aspects that influence the changes. If we define a signal-to-noise ratio here as the range of values observed across the design space divided by the natural variability, then this ratio plays an important role in the stability of the Pareto front. The results associated with responses with a small signal-to-noise ratio (the impact of changing the input values is small relative to the natural variability) are more likely change than for those responses where the signal dominates the natural variability. However, the relative priorities of the user are also important for determining if the variability of different responses is likely to be influential. Both the objective and subjective stages can show changes depending on how the natural variability changes.

Taking into account the uncertainty in the anticipated responses and their impact on the selection of a best design location to operate can be helpful for gaining confidence in future performance at that location.

## EXERCISES

**7.1** For the data in Table E7.1:

- (a) Find the Pareto front when the goal is to maximize both criteria.
- (b) Based on a scaling of the desirability functions for each of the criteria based on the entire data set (i.e., map the maximum value to 1 and the minimum

**TABLE E7.1** Data for Exercises 7.1, 7.2, and 7.3

| Obs | Criterion1 | Criterion 2 | Obs | Criterion1 | Criterion 2 |
|-----|------------|-------------|-----|------------|-------------|
| 1   | 16         | 13          | 11  | 13         | 12          |
| 2   | 11         | 12          | 12  | 18         | 15          |
| 3   | 13         | 4           | 13  | 10         | 20          |
| 4   | 14         | 5           | 14  | 15         | 11          |
| 5   | 15         | 16          | 15  | 12         | 16          |
| 6   | 17         | 3           | 16  | 15         | 12          |
| 7   | 12         | 17          | 17  | 14         | 8           |
| 8   | 13         | 12          | 18  | 17         | 6           |
| 9   | 12         | 6           | 19  | 14         | 17          |
| 10  | 11         | 12          | 20  | 14         | 4           |

value to 0) and using an additive desirability function for combining the criteria,

- (i) Find the observation that best when both criteria are weighted equally.
  - (ii) Find the observation that is best when Criterion 1 is weighted 75%.
  - (iii) Construct a mixture plot (similar to Fig. 7.3) to show which observations are best for all weight combinations of the two criteria.
  - (iv) Find the area under the Pareto front based on this scaling.
- (c) Repeat (b) with a multiplicative desirability function for combining criteria, and compare the results.
- (d) Based on a scaling of the desirability functions for each of the criteria based on only those values shown on the Pareto front (i.e., map the maximum value to 1 and the minimum value to 0) and using an additive desirability function for combining the criteria,
- (i) Find the observation that best when both criteria are weighted equally.
  - (ii) Find the observation that is best when Criterion 1 is weighted 75%.
  - (iii) Construct a mixture plot to show which observations are best for all weight combinations of the two criteria.
  - (iv) Find the area under the Pareto front based on this scaling.

**7.2** For the data in Table E7.1:

- (a) Find the Pareto front when the goal is to minimize both criteria.
- (b) Based on a scaling of the desirability functions for each of the criteria based on the entire data set and using a multiplicative desirability function for combining the criteria,
  - (i) Construct a mixture plot to show which observations are best for all weight combinations of the two criteria.
  - (ii) Find the area under the Pareto front based on this scaling.
- (c) Based on a scaling of the desirability functions for each of the criteria based on only those values shown on the Pareto front and using a multiplicative desirability function for combining the criteria,
  - (i) Construct a mixture plot to show which observations are best for all weight combinations of the two criteria.
  - (ii) Find the area under the Pareto front based on this scaling.

**7.3** For the data in Table E7.1:

- (a) Find the Pareto front when the goal is to get as close to the target value of 15 for criterion 1 and minimize criterion 2.
- (b) Based on a scaling of the desirability functions for each of the criteria based on the entire data set and using a multiplicative desirability function for combining the criteria, construct a mixture plot to show which observations are best for all weight combinations of the two criteria.
- (c) Repeat (c) using an additive desirability function and compare the results.

**7.4** For the data in Table E7.2:

- (a) Find the Pareto front when the goal is to minimize all three criteria.
- (b) Find the three pairwise Pareto fronts, and compare the number of and which observations on each of these with the front found in (a).



**TABLE E7.2 Data for Exercises 7.4**

| Obs | C1 | C2 | C3 | Obs | C1 | C2 | C3 |
|-----|----|----|----|-----|----|----|----|
| 1   | 5  | 12 | 23 | 9   | 10 | 12 | 20 |
| 2   | 2  | 17 | 28 | 10  | 3  | 17 | 26 |
| 3   | 4  | 13 | 27 | 11  | 6  | 14 | 27 |
| 4   | 5  | 19 | 25 | 12  | 4  | 19 | 24 |
| 5   | 3  | 14 | 23 | 13  | 6  | 16 | 27 |
| 6   | 7  | 10 | 24 | 14  | 8  | 18 | 22 |
| 7   | 9  | 11 | 26 | 15  | 9  | 12 | 24 |
| 8   | 2  | 18 | 30 | 16  | 5  | 11 | 30 |

- (c) Based on a scaling of the desirability functions for each of the criteria based on the entire data set and using a multiplicative desirability function for combining the criteria,
- (i) Find the observation that best when all criteria are weighted equally.
  - (ii) Find the observation that is best when the weights for the three criteria are  $(w_{C1}, w_{C2}, w_{C3}) = (0.55, 0.3, 0.15)$ .
  - (iii) Construct a table of weights that vary in increments of 0.1, find the observation that is best for each of the weight combinations.
  - (iv) Construct a trade-off plot (similar to Fig. 7.4) for the observations that are found best in (iii).
  - (v) Based on the results of (ii)–(iv), which observation would you select if the goals was to optimized for weight around  $(w_{C1}, w_{C2}, w_{C3}) = (0.55, 0.3, 0.15)$ .
- (d) Repeat (c) using an additive desirability function and compare the results.

**7.5** Theoretical results for a manufacturing process suggest that the relationships between two inputs,  $x_1$  and  $x_2$ , scaled variables in  $[-1, +1]$ , and the responses of interest are as follows:

$$\begin{aligned}\text{Yield} &= 60 + 13x_1 - 7x_2 + 5x_1x_2 \\ \text{Elasticity} &= 10 - x_1 + 2x_2 - 3x_1x_2 + x_1^2 - 2x_2^2\end{aligned}$$

The goal of the optimization is to maximize yield while attaining a target value of 11 for elasticity.

- (a) Ignoring elasticity, where is the best region to operate to maximize yield?
- (b) Ignoring yield, where is the best region to operate to obtain a response for elasticity of 11?
- (c) Find the location in the design space where the yield is maximized, conditional on the equation for elasticity specifies exactly matching the target of 11.
- (d) Find the location in the design space where the distance from the target for elasticity of 11 is minimized, subject to yield being at least 70.
- (e) By overlaying contour plots, find the region(s) in the design space where yield is greater than 70% and elasticity is between 10 and 12.

- 7.6 Using the equations for yield and elasticity from Exercise 7.5,
- For a grid of points in  $[-1, +1]$  for  $x_1$  and  $x_2$  spaced 0.1 units apart, find the Pareto front to maximize yield and get close to a target of 11 for elasticity. Plot yield and elasticity values for points on the front, and compare them to all values possible throughout the design space. How do the ranges of values compare for each response for those in the design space versus on the Pareto front?
  - With a multiplicative desirability function with the desirability score of yield scaled between  $[0,1]$  for  $[50,100]$ , and the elasticity desirability score of  $[0, 1]$  for misses from the target of  $[6, 0]$ ,
    - Find the best design location when yield and elasticity are equally weighted. What is the value of the function for each response at this location?
    - Find the best design location when yield and elasticity are weighted  $(w_{Yield}, w_{Elasticity}) = (0.75, 0.25)$
  - Repeat (d) with an additive desirability function, and desirability scores for each response based on the observed ranges from the Pareto front.
- 7.7 In their paper, Derringer and Suich (1980) illustrate their multiple-response procedure with an interesting data set. The central composite design given in Table E7.3 shows data obtained in the development of a tire tread compound on four responses: PICO Abrasion Index,  $y_1$ ; 200% modulus,  $y_2$ ; elongation at break,  $y_3$ ; hardness,  $y_4$ . Each column was taken at the 20 sets of conditions shown, where  $x_1, x_2, x_3$  are coded levels of the variables  $x_1$  = hydrated silica level,  $x_2$  = silane coupling agent level, and  $x_3$  = sulfur. The following inequalities represent desirable conditions on

TABLE E7.3 Data for Exercises 7.7 and 7.8

| Compound Number | $x_1$  | $x_2$  | $x_3$  | $y_1$ | $y_2$ | $y_3$ | $y_4$ |
|-----------------|--------|--------|--------|-------|-------|-------|-------|
| 1               | -1     | -1     | 1      | 102   | 900   | 470   | 67.5  |
| 2               | 1      | -1     | -1     | 120   | 860   | 410   | 65    |
| 3               | -1     | 1      | -1     | 117   | 800   | 570   | 77.5  |
| 4               | 1      | 1      | 1      | 198   | 2294  | 240   | 74.5  |
| 5               | -1     | -1     | -1     | 103   | 490   | 640   | 62.5  |
| 6               | 1      | -1     | 1      | 132   | 1289  | 270   | 67    |
| 7               | -1     | 1      | 1      | 132   | 1270  | 410   | 78    |
| 8               | 1      | 1      | -1     | 139   | 1090  | 380   | 70    |
| 9               | -1.633 | 0      | 0      | 102   | 770   | 590   | 76    |
| 10              | 1.633  | 0      | 0      | 154   | 1690  | 260   | 70    |
| 11              | 0      | -1.633 | 0      | 96    | 700   | 520   | 63    |
| 12              | 0      | 1.633  | 0      | 163   | 1540  | 380   | 75    |
| 13              | 0      | 0      | -1.633 | 116   | 2184  | 520   | 65    |
| 14              | 0      | 0      | 1.633  | 153   | 1784  | 290   | 71    |
| 15              | 0      | 0      | 0      | 133   | 1300  | 380   | 70    |
| 16              | 0      | 0      | 0      | 133   | 1300  | 380   | 68.5  |
| 17              | 0      | 0      | 0      | 140   | 1145  | 430   | 68    |
| 18              | 0      | 0      | 0      | 142   | 1090  | 430   | 68    |
| 19              | 0      | 0      | 0      | 145   | 1260  | 390   | 69    |
| 20              | 0      | 0      | 0      | 142   | 1344  | 390   | 70    |

the responses:

$$y_1 > 120$$

$$y_2 > 1,000$$

$$400 < y_3 < 600$$

$$60 < y_4 < 75$$

- (a) Fit appropriate response surface models for all responses.
  - (b) Develop two-dimensional contours (at  $x_3 = -1.6, -1, 0, 1, 1.6$ ) of constant response for all four responses. Can you determine sets of conditions on  $x_1, x_2$ , and  $x_3$  that meet the above requirements? If so, list them.
  - (c) Use the desirability function procedure to determine other competing conditions.
- 7.8** Reconsider the experiment in Exercise 7.7. Formulate and solve the optimization problem using:
- (a)  $y_1$ , as the objective function (maximize  $y_1$ ).
  - (b)  $y_2$  as the objective function (maximize  $y_2$ ).
  - (c) Compare the answers obtained in parts (a) and (b) with the answer from the desirability function approach in Exercise 7.7.
- 7.9** Reconsider the experiment in Experiment 7.7, with a focus on just the first two responses.
- (a) For a grid of points in  $[-1, +1]$  for  $x_1, x_2$  and  $x_3$  spaced 0.1 units apart, find the Pareto front based on two responses when the goal is to maximize  $y_1$  and  $y_2$ .
  - (b) Based on a scaling of the desirability functions for each of the criteria based on the Pareto front values and using an additive desirability function for combining the criteria,
    - (i) Find the observation that is best when both criteria are weighted equally.
    - (ii) Find the observation that is best when Criterion 1 is weighted 75%.
    - (iii) Construct a mixture plot to show which observations are best for all weight combinations of the two criteria.
  - (c) Repeat (a) and (b), where the grid of possible locations now is contained in the sphere of the design space with radius of  $\sqrt{3}$ .
- 7.10** Reconsider the experiment in Exercise 7.7.
- (a) For a grid of points in  $[-1, +1]$  for  $x_1, x_2$  and  $x_3$  spaced 0.2 units apart, find the Pareto front for the 4 responses when the goal is to maximize  $y_1$  and  $y_2$ , and minimize the distance between  $y_3$  and  $y_4$  and their respective targets of 500 and 70.
  - (b) Based on a scaling of the desirability functions for each of the criteria based on the Pareto front values and using an additive desirability function for combining the criteria,
    - (i) Find the location in the considered ranges of  $x$ 's that is best when all four criteria are weighted equally.

TABLE E7.4 Data for Exercises 7.11 and 7.12

| $x_1$ Time | $x_2$ Temperature | $x_3$ Catalyst | $y_1$ Conversion | $y_2$ Activity |
|------------|-------------------|----------------|------------------|----------------|
| -1         | -1                | -1             | 74.00            | 53.20          |
| 1          | -1                | -1             | 51.00            | 62.90          |
| -1         | 1                 | -1             | 88.00            | 53.40          |
| 1          | 1                 | -1             | 70.00            | 62.60          |
| -1         | -1                | 1              | 71.00            | 57.30          |
| 1          | -1                | 1              | 90.00            | 67.90          |
| -1         | 1                 | 1              | 66.00            | 59.80          |
| 1          | 1                 | 1              | 97.00            | 67.80          |
| -1.682     | 0                 | 0              | 76.00            | 59.10          |
| 1.682      | 0                 | 0              | 79.00            | 65.90          |
| 0          | -1.682            | 0              | 85.00            | 60.00          |
| 0          | 1.682             | 0              | 97.00            | 60.70          |
| 0          | 0                 | 1.682          | 55.00            | 57.40          |
| 0          | 0                 | 1.682          | 81.00            | 63.20          |
| 0          | 0                 | 0              | 81.00            | 59.20          |
| 0          | 0                 | 0              | 75.00            | 60.40          |
| 0          | 0                 | 0              | 76.00            | 59.10          |
| 0          | 0                 | 0              | 83.00            | 60.60          |
| 0          | 0                 | 0              | 80.00            | 60.80          |
| 0          | 0                 | 0              | 91.00            | 58.90          |

- (ii) If one experimenter is not interested in the fourth response, find the location that is best when all three of the remaining criteria are equally weighted.
- (iii) Find the location in the considered ranges of  $x$ 's that is best for

$$(w_{Y1}, w_{Y2}, w_{Y3}, w_{Y4}) = (0.7, 0.1, 0.1, 0.1)$$

- (c) Repeat (a) and (b), where the grid of possible locations now is contained in the sphere of the design space with radius of  $\sqrt{3}$ .

**7.11** Consider the chemical process experiment in Table E7.4. Find appropriate response surface models for both responses.

- (a) Use contour plots to find conditions where the conversion  $y_1$  is maximized and the thermal activity  $y_2$  is between 55 and 60.
- (b) Rework part (a) using a constrained optimization formulation.
- (c) Rework part (a) using the desirability function approach, but assume that it is necessary to make the thermal activity as close to 57.5 as possible.

**7.12** Reconsider the experiment in Exercise 7.11.

- (a) For a grid of points in the sphere of the design space with radius of  $\sqrt{3}$  for  $x_1$ ,  $x_2$  and  $x_3$  spaced 0.2 units apart, find the Pareto front for the two responses when the goal is to simultaneously maximize conversion,  $y_1$ , and minimize the distance between  $y_2$  and its target of 57.5.

- (b) Based on a scaling of the desirability functions for each of the criteria based on only those values shown on the Pareto front and using a multiplicative desirability function for combining the criteria, construct a table to show which design locations are best for all weight combinations (0, 0.1, 0.2, ...) of the two criteria.
- (c) Plot a trade-off plot of the design locations identified in (b).
- (d) Calculate the synthesized efficiency for all weight combinations in (b) for the following design locations:
- (i) Current operating conditions at  $(x_1, x_2, x_3) = (0, 0, 0)$ .
  - (ii)  $(x_1, x_2, x_3) = (1, 1, 1)$ .
  - (iii) The design location that is best when the weights for all three response desirabilities are equal [use your results from (b)].
- 7.13** Reconsider the experiment in Exercise 6.2 and Table E6.1 involving the heat transfer. Using the estimated responses and locations for maximizing thermal efficiency and transfer efficiency obtained in Exercise 6.2,
- (a) Obtain a point estimate the value of transfer efficiency when thermal efficiency is maximized.
  - (b) Obtain a 95% prediction interval for the values of the two responses when transfer efficiency is maximized.
  - (c) If the chemical engineer requires that  $y_2$  exceeds 60. Find levels (natural) of the inputs that maximize  $y_1$  subject to the constraint that  $y_2 > 60$ .
  - (d) Find the set of locations for the Pareto front if the goal is to maximize both transfer and thermal efficiency.
- 7.14** Reconsider the experiment in Exercise 6.8 and Table E6.7 involving computer networks. Using the estimated responses obtained in Exercise 6.8,
- (a) Use appropriate software to find conditions that maximize ALU, subject to the constraint that BLK is less than 0.005.
  - (b) Use the Derringer–Suich desirability function procedure to find optimal conditions when ALU cannot be less than 0.5 and BLK cannot exceed 0.10.
  - (c) For a grid of points covering the design space with scaled values spaced 0.2 units apart, find the Pareto front for the two responses when the goal is to simultaneously maximize ALU and minimize BLK.
  - (d) Based on a scaling of the desirability functions for each of the criteria based on only those values shown on the Pareto front and using a multiplicative desirability function for combining the criteria, construct a table to show which design locations are best for all weight combinations (0, 0.1, 0.2, ...) of the two criteria.
- 7.15** The method given in the text is only one possible way to implement the desirability function procedure. Below are two variations of the procedure.
1. Fit a separate response surface for each of the  $m$  responses. Create  $d_1, d_2, \dots, d_m$  and then  $D = (\prod_{l=1}^m d_l)^{1/m}$  at each design point according to values of the fitted response (the  $\hat{y}_i$ ). Then build a response surface with

$$D = f(x_1, x_2, \dots, x_k)$$

and find conditions on  $\mathbf{x}$  that maximize  $D$ . One could then make use of canonical analysis, possibly ridge analysis, and so on.

2. Compute  $d_1, d_2, \dots, d_m$  and  $D$  at each design point according to the individual observations—that is, the values  $y_1, y_2, \dots, y_m$ . Then build a response surface with the computed response  $D$ , and use appropriate methodology for finding  $\mathbf{x}$  that maximizes  $D$ .

How do you think these two methods will perform as multiple-response optimization techniques?

**7.16** Admassu and Breese (1999) performed an experiment to evaluate the ability of fishbone to sequester aqueous divalent heavy metals through absorption. The data from a  $2^4$  factorial experiment with two replicates are given Table E7.5. The levels for each of the three factors are 30 (−1) and 60 mg (+1) per 200 ml metal solution for apatite, 4 (−1) and 7 (+1) for pH. The initial lead concentration was 100 (−1) and 500 parts per million (+1). The two apatite types were fishbone (−1) and hydroxyapatite (+1). Two measurements from separate replicates for each of the responses, final lead concentration (FLC) and pH, are listed for each factor combination.

- (a) Considering only the data for the fishbone (apatite type = −1), fit a first order model with interactions for each of the responses.
- (b) Find the location in the design space where the model predicts that the final lead concentration is minimized.
- (c) Calculate the desirability score for final lead concentration if the goal is to minimize it (desirability = 1 if FLC is 0, and 0 if FLC is 2.0).
- (d) Calculate the desirability score for pH if the target value is 5.5 (desirability = 1), with values more than 2 units away being scored as 0.
- (e) For different additive desirability functions weights  $(w_{\text{FLC}}, w_{\text{pH}}) = (0,1), (0.2, 0.8), \dots, (0.8,0.2), (1,0)$ , find the best location in the design space based on the estimated responses found in (a) for fishbone.

**TABLE E7.5 Data for Exercises 7.16 and 7.17**

| Apatite | pH | Initial Lead Concentration | Apatite Type | Final Lead Concentration | Final pH   |
|---------|----|----------------------------|--------------|--------------------------|------------|
| 1       | 1  | 1                          | −1           | 1.82, 1.81               | 5.22, 5.12 |
| 1       | 1  | −1                         | −1           | 0.01, 0.00               | 6.84, 6.61 |
| 1       | −1 | 1                          | −1           | 1.11, 1.04               | 3.35, 3.34 |
| 1       | −1 | −1                         | −1           | 0.00, 0.01               | 5.77, 6.25 |
| −1      | 1  | 1                          | −1           | 2.11, 2.18               | 5.29, 5.06 |
| −1      | 1  | −1                         | −1           | 0.03, 0.05               | 5.93, 6.02 |
| −1      | −1 | 1                          | −1           | 1.70, 1.69               | 3.39, 3.34 |
| −1      | −1 | −1                         | −1           | 0.05, 0.05               | 4.50, 4.74 |
| 1       | 1  | 1                          | 1            | 0.11, 0.12               | 3.49, 3.46 |
| 1       | 1  | −1                         | 1            | 0.00, 0.00               | 5.84, 5.90 |
| 1       | −1 | 1                          | 1            | 0.80, 0.76               | 2.70, 2.74 |
| 1       | −1 | −1                         | 1            | 0.03, 0.05               | 3.36, 3.24 |
| −1      | 1  | 1                          | 1            | 1.03, 1.05               | 3.22, 3.22 |
| −1      | 1  | −1                         | 1            | 0.00, 0.00               | 5.53, 5.43 |
| −1      | −1 | 1                          | 1            | 1.34, 1.26               | 2.81, 2.79 |
| −1      | −1 | −1                         | 1            | 0.06, 0.07               | 3.28, 3.28 |

- 7.17** Reconsider the experiment in Exercise 7.16 and Table E7.5 involving the apatite data.
- Repeat the analysis in Exercise 7.16 for the data for the hydroxyapatite (apatite type = 1).
  - Which of the apatite types is more effective at reducing the final lead concentration?
  - Comment on the similarity or difference in the recommended combination of factors between the two types of apatites for the different weight combinations of the desirability function explored.
- 7.18** Modi et al. (2001) describe an experiment to compare the wear rate of two materials, a zinc-aluminum alloy and a composite. Three different factors affecting wear were considered: the slide distance ( $x_1 = -1 \rightarrow 25\text{m}, +1 \rightarrow 125\text{m}$ ), applied load ( $x_2 = -1 \rightarrow 1\text{N}, +1 \rightarrow 7\text{N}$ ), and abrasive size ( $x_3 = -1 \rightarrow 23\mu\text{m}, +1 \rightarrow 275\mu\text{m}$ ).
- Fit a model of the form
 
$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3 + \varepsilon$$
 for the wear rate of each material.
  - For the zinc-aluminum alloy model, find the set of contours where the wear rate is predicted to be 0.1.
  - For the composite material, find the set of contours where the wear rate is predicted to be 0.1.
  - Find the location in the design space that simultaneously provides a wear rate as close to 0.1 for both materials as possible. *Note:* Consider how to define a measure that balances the performance of both materials.
- 7.19** Reconsider the experiment in Exercise 7.18 and Table E7.6 involving the abrasive wear study.
- Using a 95% confidence interval based on the models estimated in Exercise 7.18, find the regions in the design space where the estimated model might be expected to produce a value for the wear rate of 0.1.
  - Repeat (a) using a 95% prediction interval.
  - Based on the results from (a) and (b) for the zinc-aluminum alloy and the composite material, discuss how likely it is to have to two materials simultaneously give similar results for a wear rate of 0.1.

**TABLE E7.6** Data for Exercises 7.18 and 7.19

| Sliding Distance | Applied Load | Abrasive Size | Zinc Wear Rate | Composite Wear Rate |
|------------------|--------------|---------------|----------------|---------------------|
| 1                | 1            | 1             | 0.192          | 0.112               |
| 1                | 1            | -1            | 0.133          | 0.054               |
| 1                | -1           | -1            | 0.044          | 0.006               |
| -1               | -1           | -1            | 0.049          | 0.011               |
| -1               | 1            | 1             | 0.358          | 0.275               |
| -1               | -1           | 1             | 0.041          | 0.018               |
| 1                | -1           | 1             | 0.03           | 0.006               |
| -1               | 1            | -1            | 0.22           | 0.0099              |

**TABLE E7.7 Data for Exercises 7.20 and 7.21**

| $X_1$ | $X_2$ | $X_3$ | $X_4$ | $Y_1$ | $Y_2$ | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $Y_1$ | $Y_2$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1     | 250   | 60    | 1     | 2.58  | 0.24  | 2     | 275   | 73    | 3     | 1.33  | 0.49  |
| 1     | 250   | 60    | 2     | 1.32  | 0.26  | 2     | 275   | 85    | 1     | 3.34  | 0.25  |
| 1     | 250   | 60    | 3     | 0.61  | 0.5   | 2     | 275   | 85    | 2     | 2.26  | 0.25  |
| 1     | 250   | 73    | 1     | 2.56  | 0.25  | 2     | 275   | 85    | 3     | 1.44  | 0.27  |
| 1     | 250   | 73    | 2     | 1.37  | 0.25  | 2     | 300   | 60    | 1     | 3.88  | 0.24  |
| 1     | 250   | 73    | 3     | 0.74  | 0.31  | 2     | 300   | 60    | 2     | 2.71  | 0.31  |
| 1     | 250   | 85    | 1     | 2.75  | 0.25  | 2     | 300   | 60    | 3     | 1.82  | 0.53  |
| 1     | 250   | 85    | 2     | 1.72  | 0.25  | 2     | 300   | 73    | 1     | 3.89  | 0.25  |
| 1     | 250   | 85    | 3     | 0.98  | 0.25  | 2     | 300   | 73    | 2     | 2.68  | 0.26  |
| 1     | 275   | 60    | 1     | 2.8   | 0.24  | 2     | 300   | 73    | 3     | 1.81  | 0.38  |
| 1     | 275   | 60    | 2     | 1.89  | 0.26  | 2     | 300   | 85    | 1     | 3.88  | 0.25  |
| 1     | 275   | 60    | 3     | 1.3   | 0.42  | 2     | 300   | 85    | 2     | 2.46  | 0.25  |
| 1     | 275   | 73    | 1     | 3.04  | 0.25  | 2     | 300   | 85    | 3     | 1.72  | 0.26  |
| 1     | 275   | 73    | 2     | 2.09  | 0.25  | 3     | 250   | 60    | 1     | 1.8   | 0.93  |
| 1     | 275   | 73    | 3     | 1.55  | 0.27  | 3     | 250   | 60    | 2     | 0.78  | 1.57  |
| 1     | 275   | 85    | 1     | 3.35  | 0.25  | 3     | 250   | 60    | 3     | 0.2   | 1.81  |
| 1     | 275   | 85    | 2     | 2.41  | 0.25  | 3     | 250   | 73    | 1     | 2.02  | 0.26  |
| 1     | 275   | 85    | 3     | 1.9   | 0.25  | 3     | 250   | 73    | 2     | 0.94  | 0.93  |
| 1     | 300   | 60    | 1     | 3.87  | 0.24  | 3     | 250   | 73    | 3     | 0.31  | 1.39  |
| 1     | 300   | 60    | 2     | 2.84  | 0.25  | 3     | 250   | 85    | 1     | 2.45  | 0.25  |
| 1     | 300   | 60    | 3     | 2.38  | 0.3   | 3     | 250   | 85    | 2     | 1.36  | 0.27  |
| 1     | 300   | 73    | 1     | 4.26  | 0.25  | 3     | 250   | 85    | 3     | 0.66  | 0.6   |
| 1     | 300   | 73    | 2     | 3.18  | 0.25  | 3     | 275   | 60    | 1     | 2.7   | 0.45  |
| 1     | 300   | 73    | 3     | 2.7   | 0.26  | 3     | 275   | 60    | 2     | 1.5   | 1.1   |
| 1     | 300   | 85    | 1     | 4.64  | 0.25  | 3     | 275   | 60    | 3     | 0.83  | 1.34  |
| 1     | 300   | 85    | 2     | 3.5   | 0.25  | 3     | 275   | 73    | 1     | 3.06  | 0.26  |
| 1     | 300   | 85    | 3     | 3     | 0.25  | 3     | 275   | 73    | 2     | 1.91  | 0.56  |
| 2     | 250   | 60    | 1     | 2.51  | 0.25  | 3     | 275   | 73    | 3     | 1.16  | 1.01  |
| 2     | 250   | 60    | 2     | 1.32  | 0.51  | 3     | 275   | 85    | 1     | 3.4   | 0.25  |
| 2     | 250   | 60    | 3     | 0.6   | 0.94  | 3     | 275   | 85    | 2     | 2.33  | 0.25  |
| 2     | 250   | 73    | 1     | 2.65  | 0.25  | 3     | 275   | 85    | 3     | 1.47  | 0.36  |
| 2     | 250   | 73    | 2     | 1.47  | 0.3   | 3     | 300   | 60    | 1     | 3.83  | 0.36  |
| 2     | 250   | 73    | 3     | 0.75  | 0.58  | 3     | 300   | 60    | 2     | 2.5   | 0.78  |
| 2     | 250   | 85    | 1     | 2.71  | 0.25  | 3     | 300   | 60    | 3     | 1.65  | 0.98  |
| 2     | 250   | 85    | 2     | 1.63  | 0.25  | 3     | 300   | 73    | 1     | 4.17  | 0.25  |
| 2     | 250   | 85    | 3     | 0.86  | 0.29  | 3     | 300   | 73    | 2     | 2.85  | 0.34  |
| 2     | 275   | 60    | 1     | 3.39  | 0.24  | 3     | 300   | 73    | 3     | 1.97  | 0.69  |
| 2     | 275   | 60    | 2     | 2.2   | 0.43  | 3     | 300   | 85    | 1     | 4.47  | 0.25  |
| 2     | 275   | 60    | 3     | 1.31  | 0.68  | 3     | 300   | 85    | 2     | 3.22  | 0.25  |
| 2     | 275   | 73    | 1     | 3.29  | 0.24  | 3     | 300   | 85    | 3     | 2.36  | 0.3   |
| 2     | 275   | 73    | 2     | 2.14  | 0.27  |       |       |       |       |       |       |

(d) Within the design space, for what value of wear rate are the two materials most similar? Where in the design space does this value occur? Does this result seem to match your intuition?

**7.20** Barghash and Alkaabneh (2014) describe an experiment involving four factors each considered at three levels to model the shrinkage ( $Y_1$ ) and warpage ( $Y_2$ ) in millimeters of an injection molding process. Table E7.7 shows the data from a  $3^4 =$



81-run experiment using a full factorial design. The four factors considered are as follows:

|  | -1  | 0    | +1  |
|--|-----|------|-----|
| $X_1$ : Filling time (sec)                         | 1   | 2    | 3   |
| $X_2$ : Melting temperature ( $^{\circ}\text{C}$ ) | 250 | 275  | 300 |
| $X_3$ : Mold temperature ( $^{\circ}\text{C}$ )    | 60  | 72.5 | 85  |
| $X_4$ : Pressure holding time (sec)                | 1   | 2    | 3   |

- Fit a second-order model to  $Y_1$  and find the location in the design space that minimizes the amount of shrinkage of the injection molding process.
- Fit a second-order model to  $Y_2$  and find the location in the design space that minimizes the amount of warpage.
- Weighting the two responses equally and scaling the additive desirability function with values of 0 and 1 for the worst and best values observed in the design space, respectively, find the optimal combination of the factors to balance minimizing the shrinkage and warpage.
- If the experimenters are willing to accept shrinkage of at most 2 mm, what is the best estimated warpage that can be achieved. What combination of factors leads to this result?
- If the experimenters are willing to accept warpage of at most 0.3 mm, what is the best estimated shrinkage that can be achieved. What combination of factors leads to this result?

**7.21** Reconsider the experiment in Exercise 7.20 and Table E7.7 involving an injection molding process.

- Construct the Pareto front of best values to balance minimizing shrinkage and warpage across the design space.
- Using the results from (a), find the set of solutions that are best for at least one weight combination of an additive desirability function based on the range of shrinkage and warpage values on the Pareto front.
- Repeat (a) with a multiplicative desirability function.
- If the experimenters wish to approximately balance the shrinkage and warpage for the chosen solution, what factor combination would you recommend? Explain why you think this is a preferred solution.
- Repeat (d) if the scaling selected for each factor in the desirability function is based on the best and worst values observed throughout the design space.

### Short Answer Questions

**7.22** Answer the following short answer questions:

- Why is it helpful to divide the process of multiple response/criteria optimization into objective and subjective stages?
- Describe some of the aspects that make comparing different criteria difficult.

- (c) What are some of the limitations of looking at contour plots of all of the responses to find a good multiple response solution.
- (d) Give a situation when the threshold approach might be advantageous.
- (e) Compare how the additive and multiplicative desirability functions penalize poor results for one or more of the responses.
- (f) Give a situation where two experimenters might have different priorities for their choice of how to weight the importance of the responses.
- (g) Justify why the Pareto front approach allows you to exclude some of the possible solutions as not logical choices for optimal.
- (h) Why is it important to consider the uncertainty in estimating the responses when selecting the optimal solution for a multiple response optimization problem?

**7.23** Answer the following True/False questions:

- (a) The desirability function converts the response variable into a new variable that varies over the range  $-1$  to  $+1$ .
- (b) A solution on the Pareto front is best for some combination of weights using an additive desirability function.
- (c) An optimal solution found using the threshold approach is also located on the Pareto front.
- (d) The solution found using the desirability function approach is robust to the choice of scaling for each of the criteria.
- (e) Using an additive desirability function penalizes poor values of the criterion more than using a multiplicative desirability function with the same scaling for each criteria.
- (f) The amount of trade-off between responses can be manipulated through the choice of design.
- (g) Creating a grid of locations in the design space over which to evaluate the estimated response surfaces is an effective way to find factor combinations to optimize multiple responses.
- (h) The synthesized efficiency plot is a graphical summary that allows comparisons between different potential solutions.
- (i) The smaller area under the Pareto front indicates a more severe trade-off between two criteria.
- (j) Using the worst-case boundary of a prediction interval for new observations is an optimization approach to consider the uncertainty of the estimated response surfaces of the responses.

---

# 8

---

## DESIGN OF EXPERIMENTS FOR FITTING RESPONSE SURFACES—I

Many of our examples in Chapter 6 displayed the fitting of second-order response surfaces with data taken from real-life applications of response surface methodology (RSM). In those examples, the chosen experiment design possesses properties required to fit a second-order response surface—that is:

1. At least three levels of each design variable.
2. At least  $1 + 2k + k(k - 1)/2$  distinct design points.

The above represent minimum conditions for fitting a second-order model. The reader has also noticed the reference to the central composite design in Chapter 6. This second-order design was the design of choice for almost all of the illustrations. The central composite design is the most popular second-order design in use by practitioners. As a result, it will receive considerable attention in the sections that follow. First, however, some general discussion is presented regarding the choice of designs for response surface analysis.

### 8.1 DESIRABLE PROPERTIES OF RESPONSE SURFACE DESIGNS

There are many classes of designed experiments in the literature, and there are many criteria on which designs are based. Indeed, there are many computer packages that give optimal designs based on special criteria and input from the user. Special design criteria and important issues associated with computer-generated design of experiments will be discussed in a later section and in Chapter 9. However, it is important for the reader to first

review a set of properties that should be taken into account when the choice of a response surface design is made. Some of the important characteristics are as follows:

1. Result in good fit of the model to the data.
2. Provide good model parameter estimates.
3. Provide a good distribution of prediction variance of the response,  $\text{Var} [\hat{y}(x)] / \sigma^2$ , throughout the region of interest.
4. Provide an estimate of “pure” experimental error.
5. Give sufficient information to allow for lack of fit test.
6. Provide a check on the homogeneous variance assumption.
7. Be insensitive (robust) to the presence of outliers in the data.
8. Be robust to errors in the control of design levels.
9. Allow models of increasing order to be constructed sequentially.
10. Allow for experiments to be done in blocks.
11. Be cost-effective.

As one can readily see, not all of the above properties are required in every RSM application. However, most of them should be given serious consideration on each occasion in which one designs experiments. Most of the properties are self-explanatory. To help organize our thinking about the characteristics on this list, we can divide the list into several categories. Items 1–4 assume that the practitioner has made good assumptions about the nature of the underlying relationship between the inputs and the response; and under the assumption that the model is correct, the goal is to have good estimation of model parameters and prediction of new observations when we use the model. The reader has been exposed to the notion of prediction variance in Chapters 2, 3, and 6. In later sections we shall discuss the importance of stability of prediction variance. This is often the most common category on which emphasis is focused when selecting a designed experiment.

Items 5 and 6 seek to provide ways of evaluating the assumptions of the model. All models have associated assumptions, and it is helpful if the data collected in the experiment can allow some feedback about the suitability of these assumptions to process under consideration. Items 7 and 8 focus on how the experiment will be affected if something goes wrong. The goal is to plan an experiment that can withstand some less-than-ideal outcome and still generate useful results. Item 7 is particularly important if one expects outliers to occur. Items 9 and 10 seek to allow flexible implementation of the experiment and leveraging of the results as part of the sequential nature of many experimental learning cycles. Finally, item 11 reminds us that there are often cost constraints imposed on experiments. Larger experiments often can provide improved characteristics of the first 10 items, but at the expense of driving up the total cost of the experiment and preventing those resources from being used for other purposes.

There are several goals behind the introduction of the list of 11 characteristics at this point in the text. Of primary importance is a reminder to the reader that designing an experiment is not necessarily easy and should involve balancing multiple objectives, not just focusing on a single characteristic. Similar to the discussions in Chapter 7 where optimization of a product or process involved considerations of multiple aspects, designing an experiment usually involves balancing multiple objectives. Indeed, several of the 11 items may be important and yet the researcher may not be completely aware of their relative importance.

Some items do conflict with each other. As a result, there are trade-offs that almost always exist when one chooses an appropriate design. For example, good choices for where to allocate runs under the assumption that the model selected *is* correct differ considerably for where we would select runs if we are trying to evaluate *if* the model is correct. Similarly, assuming that the model is correct versus wanting to protect against model misspecification would lead to different choices of designs. Another potential trade-off involves whether we think that the implementation of the experiment will run smoothly or whether there might be some complications. Finally, cost trades off with most of the other categories: A larger experiment can help improve the results for most of the first four categories, but at the cost of using additional resources. An excellent discourse on the many considerations that one must consider in choosing a response surface design is given by Box and Draper (1975).

## 8.2 OPERABILITY REGION, REGION OF INTEREST, AND METRICS FOR DESIRABLE PROPERTIES

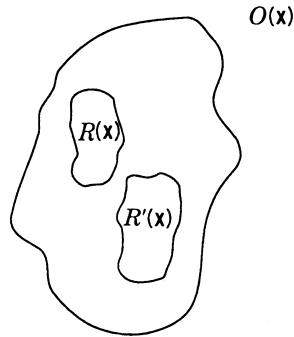
At the end of Chapter 6 we discussed the choice of ranges on the design variables; this naturally leads to an introduction of the notion of the region of interest in the design variables. As we indicated in that chapter, we are assuming that the true function relating  $y$  to the design variables is unknown. In fact,

$$E(y) = f(\mathbf{x}, \theta)$$

and we are assuming that in the region of interest,  $R(\mathbf{x})$ ,  $f$  is well approximated by a low-order polynomial. This, of course, introduced the notion of the first-order or second-order response functions. They are approximations (justified by a Taylor series expansion of  $f$ ) in the confined region  $R$ .

It is important for the reader to understand that while the experimental design may be confined to  $R$ , the region  $R$  may change from experiment to experiment (say, for example, in a steepest ascent experiment, as described in Chapter 5). However, there is a secondary region called the **region of operability**,  $O(\mathbf{x})$ . This is the region in which the equipment, electronic system, chemical process, drug, or the like, works, and it is theoretically possible to do the experiment and observe response values. In some cases, data points in the design may be taken outside  $R(\mathbf{x})$ . Obviously, if we take data too far outside  $R(\mathbf{x})$ , then the adequacy of our response surface representation (polynomial model) may be in question. Though the framework of the regions of interest and operability is important for the reader to understand, assuming clear and specific knowledge of  $R(\mathbf{x})$  or  $O(\mathbf{x})$  in a given situation may be too idealistic. The region  $R$  changes, and at times  $O$  is not truly known until much is known about the process.

Figure 8.1 gives the reader a pictorial depiction of  $O$  and  $R$ . In this case, with two different experiments we have  $R(\mathbf{x})$  and  $R'(\mathbf{x})$ , both being considerably more confined than  $O$ . The region  $R$  is the current best guess at where the optimum is. It is also the region in which we feel that  $\hat{y}(\mathbf{x})$  should be a good predictor or, say, a good estimator of  $f(\mathbf{x}, \theta)$ . However, the user should constantly be aware of potential model inadequacy in cases where  $\hat{y}(\mathbf{x})$  is not a sufficiently good approximation of  $f(\mathbf{x}, \theta)$ .



**Figure 8.1** Region of operability and region of interest.

### 8.2.1 Metrics for Desirable Properties

The list of different criteria useful for assessing the “goodness” of a designed experiment identified in Section 8.1 become much more useful when we associate them with quantitative metrics which allow direct comparison between alternative designs. In this section we focus on metrics associated with the first category of design properties, which assume that the model appropriately reflects the underlying relationship between input variables and the response. Our goal with this category of metrics is to capture the adequacy of the design to estimate model parameters and predict the response value for new observations. In Section 3.6 of Chapter 3, the concept of optimal designs was introduced. Recall that maximizing the determinant of  $\mathbf{X}'\mathbf{X}$  (where  $\mathbf{X}$  is the model matrix introduced in Section 2.2) gives a design that provides estimates of the model parameters,  $\beta$  (for the model  $\mathbf{y} = \mathbf{X}\beta + \epsilon$ ) with the smallest possible variance. Known as *D-optimality*, this criterion focuses on the second item in the list in Section 8.1, namely good model parameter estimates.

A key ingredient in choosing a design that minimizes the variance of the model parameters involves examining **orthogonality**. The terms **orthogonal design**, **orthogonal array**, and **orthogonal main effect plan** received considerable attention in the 1950s and 1960s. An orthogonal design is one for which  $\mathbf{X}'\mathbf{X}$  is a diagonal matrix. The above definition implies, of course, that the columns of  $\mathbf{X}$  are mutually orthogonal. It should be clear that if two columns are orthogonal, the levels of the two corresponding variables are linearly independent. The implication is that the roles of the two terms in the model can be being assessed independently of each other. This underscores the virtues of orthogonality. As we examine particular designed experiments for various models in subsequent sections of this chapter, we will examine their orthogonality properties.

Alternately, other optimality criteria, such as *G-* and *I-*optimality (also introduced in Section 3.6), focus on the third item in the list of desirable properties, a good distribution of predicted variance of the response throughout the region of interest. With a small predicted variance value at a given combination of input values, we improve our ability to precisely predict the value of the response at that location in the design space. Recall that unscaled prediction variance is defined as

$$UPV = \text{Var} [\hat{y}(\mathbf{x})] / \sigma^2$$

and can be evaluated at any location in the design space. JMP uses Fraction of Design Space (FDS) plots to show a summary of the distribution of the unscaled prediction variance

throughout the design region. A *G-optimal design* is one that minimizes the worst case, or maximum prediction variance throughout the region of interest. Often the worst prediction occurs near the edges of the selected design space (or occasionally at the center), and so evaluating G-optimality requires exploring prediction variance values throughout the design space, and particularly at the edges. Since the experimenter often does not know where the prediction of new observations will be needed, this criterion focuses on making sure that the least desirable prediction variance is not too bad.

Alternately, *I-optimality* focuses on the average prediction variance in the design space, and it seeks to make that as small as possible. In statistics applications, we are more often focused on the center of the distribution of values, and so I-optimality examines the average prediction variance across all possible locations in the space where new observations might be obtained. In Chapter 9, more details will be given about these criteria as well as how they can be incorporated into optimization algorithms to generate tailored designs for a specific experiment.

Also related to the characteristics of the prediction variance throughout the design space is the concept of **rotatability**. Introduced in Section 2.8, the concept of rotatability focuses on the symmetry of prediction variance in the region of interest. A rotatable design is one that has constant prediction variance values for all locations at equal distance from the center of the design space. For example, a rotatable design would have the same prediction variance value for any new observation in the design space that was 1 unit away from the origin, when examining the scaled ranges of the input variables.

### 8.2.2 Model Inadequacy and Model Bias

In this section, we provide details concerning the impact of specifying a response surface model incorrectly. This is particularly important since one is almost always making use of empirical models. Thus, in many instances, more than one empirical model may be reasonable and some are better than others.

Consider a situation in which we assume a model of the form  $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}$ . After the data are collected and fit, we have a vector of fitted values,

$$\hat{\mathbf{y}} = \mathbf{X}_1\mathbf{b}_1 \quad (8.1)$$

where

$$\mathbf{b}_1 = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{y}$$

Here we implicitly make the assumption that

$$E(\mathbf{y}) = \mathbf{X}_1\boldsymbol{\beta}_1$$

However, there may be a better approximation of  $E(\mathbf{y})$  given by

$$E(\mathbf{y}) = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 \quad (8.2)$$

where, say,  $\mathbf{X}_1$  is an  $N \times p_1$  matrix, and  $\mathbf{X}_2$  is  $N \times p_2$ . The additional term in Equation 8.2,  $\mathbf{X}_2\boldsymbol{\beta}_2$  contains additional linear model terms which provide more flexibility in the modeling

of the response. Examples of this would be higher order terms or additional interaction terms. Then the expected value of  $\mathbf{b}_1$  is given by

$$\begin{aligned} E(\mathbf{b}_1) &= (\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1E[\mathbf{y}] \\ &= (\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1E[\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}] \\ &= (\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1\mathbf{X}_1\boldsymbol{\beta}_1 + (\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1\mathbf{X}_2\boldsymbol{\beta}_2 \\ &= \boldsymbol{\beta}_1 + (\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1\mathbf{X}_2\boldsymbol{\beta}_2 \\ &= \boldsymbol{\beta}_1 + \mathbf{A}\boldsymbol{\beta}_2 \end{aligned}$$

Here  $\boldsymbol{\varepsilon}$  is the usual random error, and  $\mathbf{A} = (\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1\mathbf{X}_2$  is called the **alias matrix** (introduced in Section 4.7). The alias matrix transmits bias errors to the estimate  $\mathbf{b}_1$ . Understanding the effect of bias on the estimates of the model is applicable for many types of models. There is considerable direct application to first order models, such as fractional factorials discussed in Chapter 4 (see Exercises 8.7, 8.8, and 8.9).

As one might expect, if there is appreciable model misspecification, bias errors also appear in the fitted values. For, in the above scenario, the expected value of the vector of fitted values is given by

$$\begin{aligned} E(\hat{\mathbf{y}}) &= \mathbf{X}_1E(\mathbf{b}_1) \\ &= \mathbf{X}_1(\boldsymbol{\beta}_1 + \mathbf{A}\boldsymbol{\beta}_2) \\ &= \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_1\mathbf{A}\boldsymbol{\beta}_2 \end{aligned}$$

and thus

$$\begin{aligned} E(\hat{\mathbf{y}}) - E(\mathbf{y}) &= [\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_1\mathbf{A}\boldsymbol{\beta}_2] - [\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2] \\ &= (\mathbf{X}_1\mathbf{A} - \mathbf{X}_2)\boldsymbol{\beta}_2 \end{aligned} \tag{8.3}$$

When the true model is best approximated by Equation 8.2, bias is transmitted to both model coefficients and fitted values. Equation 8.3 is important in that it plays a role in the lack-of-fit test described in Chapter 6. If the model is misspecified as we have described here, the residual mean square,  $s^2$ , in the analysis of variance is also biased. Instead of estimating  $\sigma^2$  as expected, it has expectation

$$E(s_2) = \sigma^2 + \frac{1}{p_1} \left\{ \boldsymbol{\beta}'_2 [\mathbf{X}_1\mathbf{A} - \mathbf{X}_2]' [\mathbf{X}_1\mathbf{A} - \mathbf{X}_2] \boldsymbol{\beta}_2 \right\}$$

The bias in  $s^2$  described above is intuitive. It is simply proportional to the sum of squares of the biases in the fitted values. This is reasonable since the lower mean square is itself proportional to the sum of squares of deviations between  $y$  observations and fitted values. If we call the bias in  $s^2$

$$\frac{1}{p_1} \left\{ \boldsymbol{\beta}'_2 \mathbf{R}' \mathbf{R} \boldsymbol{\beta}_2 \right\} \tag{8.4}$$

where  $\mathbf{R} = \mathbf{X}_1\mathbf{A} - \mathbf{X}_2$ , then it becomes clear that the lack-of-fit test, when significant, is detecting the existence of a nonzero  $\boldsymbol{\beta}_2$  through the quadratic form  $\boldsymbol{\beta}'_2 \mathbf{R}' \mathbf{R} \boldsymbol{\beta}_2$ .



In addition to lack of fit, the vector of biases in fitted value given in Equation 8.3 may play an important role in the choice of an experimental design. One point of view is to determine the design based on control of two types of error, **variance error** and **bias error**. Thus, we have a loss function type criterion. For each observation,  $y_j$ , we want to make small the loss

$$\begin{aligned} E [\hat{y}_j - f(\mathbf{x}_j, \boldsymbol{\beta})]^2 &= E [\hat{y}_j - E(\hat{y}_j) + E(\hat{y}_j) - f(\mathbf{x}_j, \boldsymbol{\beta})]^2 \\ &= E [\hat{y}_j - E(\hat{y}_j)]^2 + [E(\hat{y}_j) - f(\mathbf{x}_j, \boldsymbol{\beta})]^2 \\ &= \text{Var}(\hat{y}_j) + [\text{Bias}(\hat{y}_j)]^2 \end{aligned} \quad (8.5)$$

The term  $\text{Var}(\hat{y}_j)$  is clear. It has been discussed in Chapter 6 and will receive further attention in this chapter. The term  $[\text{Bias}(\hat{y}_j)]^2$  is the squared bias in the fitted values if indeed  $E(\mathbf{y}) \neq \mathbf{X}_1 \boldsymbol{\beta}_1$ . The bias term is related to Equation 8.4 as

$$\frac{1}{\rho_1} \{ \boldsymbol{\beta}'_2 \mathbf{R}' \mathbf{R} \boldsymbol{\beta}_2 \} = \sum_{j=1}^N [\text{Bias}(\hat{y}_j)]^2$$

The statistics literature contains a considerable amount of information on experimental design for regression or response surface modeling. However, the majority of this technical or theoretical information deals with variance-oriented design criteria—that is, the type of design criteria that ignores model misspecification. In other words, these criteria essentially assume that the model specified by the user is correct. In what follows, we shall initially focus on variance-type criteria. More recently, work has been considered where designs are evaluated not only by assessing the variance error, but also considering potential bias error. Because the bias term contains  $\boldsymbol{\beta}_2$ , some additional assumptions for the magnitude of these missing terms is needed to quantify the size of the bias. See Draper and Sanders (1988), Vining and Myers (1991), Allen et al. (2003), and Anderson-Cook et al. (2009a) for a variety of approaches to considering bias error. However, the errors associated with model misspecification will be revisited in Chapter 10. In what follows, we begin with experimental designs for first-order response models.

### 8.3 DESIGN OF EXPERIMENTS FOR FIRST-ORDER MODELS AND FIRST-ORDER MODELS WITH INTERACTIONS

One of the primary goals of running a designed experiment aimed at fitting a first-order model, or a first-order model with interactions, is to explore which of the factors substantially impact the values of the response. Often called *screening designs*, these experiments commonly focus on good estimation of the model parameters and providing a ranking of how much the different inputs change the response. The name “screening” implies that the experimenter is looking to divide the inputs into groups: those that definitely impact the response, those that do not appear to have a meaningful impact on the response over the range of input values explored, and potentially a third category where there is some ambiguity over whether the contribution of that factor is important or not.

This understanding of the contributions of the different inputs, through main effects and potentially through two-factor interactions, allows the experimenter to prioritize where to focus for subsequent experiments that seek to optimize the process or product. Reducing a large number of potential input factors down to a more manageable few for further exploration is a key step in many sequential design of experiments progressions.

Consider a situation in which  $N$  experimental runs are conducted on  $k$  design variables  $x_1, x_2, \dots, x_k$  and a single response  $y$ . A model is postulated of the type

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i \quad (i = 1, 2, \dots, N)$$

As a result, a fitted model is given by

$$\hat{y}_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_k x_{ik}$$

Here, of course, the  $b_j$  are found by the method of least squares. An intuitive variance-based criterion if we are interested in good estimation of the model parameters involves the variance of the regression coefficients—that is,  $\text{Var}(b_j)$  ( $j = 1, 2, \dots, k$ ). Let us assume that each design variable is in coded design units and that the ranges for the design levels are such that  $x_j \in [-1, +1]$ .

### 8.3.1 The First-Order Orthogonal Design

In Section 8.2, the idea of orthogonality was introduced as a desirable property to incorporate into a designed experiment, since it allowed the separate assessment of the role of two variables. In the first-order situation, an orthogonal design in which the ranges on these variables are set at the edges of the range results in minimizing  $\text{Var}(b_j)$  on a per observation basis. In other words,

For the first-order model and a fixed sample  $N$ , if  $X_j \in [-1, +1]$  for  $j = 1, 2, \dots, k$ , then  $\text{Var}(b_i)/\sigma^2$  for  $i = 1, 2, \dots, k$  is minimized if the design is orthogonal and all  $x_i$  levels in the design are  $\pm 1$  for  $i = 1, 2, \dots, k$ .

Thus, the elements on the diagonals of  $(\mathbf{X}'\mathbf{X})^{-1}$  are minimized [recall that  $\text{Var}(\mathbf{b}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$ ] by making off-diagonals of  $\mathbf{X}'\mathbf{X}$  zero and by forcing the diagonals of  $\mathbf{X}'\mathbf{X}$  to be as large as possible. The notion of **linear independence** and **variable levels at  $\pm 1$  extremes** as a desirable set of conditions should certainly be intuitive to the reader.

**Example 8.1 Variance-Optimal First-Order Designs I** Two-level factorial plans and fractions of resolution III and higher do, in fact, minimize the variances of all coefficients (variances scaled by  $\sigma^2$ ) on a per observation basis. To illustrate this, consider the following

$2^{4-1}$  fractional factorial (resolution IV) with the defining relation  $I=ABCD$ . The matrix  $\mathbf{X}$  is given by

$$\mathbf{X} = \begin{array}{c} \begin{array}{cccc} & x_1 & x_2 & x_3 & x_4 \end{array} \\ \left[ \begin{array}{ccccc} 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{array} \right] \end{array}$$

One can easily determine that  $\mathbf{X}'\mathbf{X} = 8\mathbf{I}_5$  and  $(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{8}\mathbf{I}_5$ . Indeed, one can say that *no other design with eight experimental runs in this region can result in variances smaller than  $\sigma^2/8$* . It should be apparent to the reader that a full  $2^4$  factorial results in all coefficient variances being  $\sigma^2/16$ . In other words, if the size of the design is doubled, the variances of coefficients are cut in half. However, both are considered optimal designs on a per observation basis.

**Example 8.2 Variance-Optimal First-Order Designs II** Consider a situation in which there are three design variables and the user wishes to use eight experimental runs but also to have design replication. A  $2^{3-1}$  fractional factorial (resolution III) is used with each design point replicated. The matrix  $\mathbf{X}$  for the design is as follows (defining relation  $I=ABC$ ):

$$\mathbf{X} = \begin{array}{c} \begin{array}{ccc} & x_1 & x_2 & x_3 \end{array} \\ \left[ \begin{array}{cccc} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{array} \right] \end{array}$$

Note that the columns are orthogonal and all levels are at  $\pm 1$  extremes. Thus the design is **variance optimal** for the model

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3$$

That is, all coefficients in the above model have minimum variance over all designs with sample size  $N=8$ . In fact, the variance-covariance matrix is given by

$$\text{Var}(\mathbf{b}) = (\sigma^2/8)\mathbf{I}_4$$

It is interesting that for this situation a full  $2^3$  factorial results in the same variance-covariance matrix. Though the two designs are quite different, from a variance point of view

they are equivalent for the first-order model in three design variables. Obviously, in the case of the  $2^{3-1}$  fraction (replicated) there are no degrees of freedom for lack of fit, whereas the  $2^3$  factorial enjoys four lack-of-fit degrees of freedom that are attributable to  $x_1x_2$ ,  $x_1x_3$ ,  $x_2x_3$ , and  $x_1x_2x_3$ . On the other hand, the replicated one-half fraction allows four degrees of freedom for replication error (pure error), and, of course, the full  $2^3$  possesses no degrees of freedom for replication error. As a result, even though they are variance-equivalent, the two orthogonal designs would find use in somewhat different circumstances.

It should be apparent to the reader that for the use of two-level factorials or fractions for a first-order model, orthogonality (and hence variance optimality) is obtained with resolution at least III. In the case of resolution III, the  $\mathbf{x}_i$  columns in  $\mathbf{X}$  will be aliased with the  $\mathbf{x}_j\mathbf{x}_k$  columns. However, no  $\mathbf{x}_i$  column will be aliased with any  $\mathbf{x}_j$  column. In the case of  $2^k$  factorials or *regular fractions*, *any two columns in  $\mathbf{X}$  that are not aliased are, indeed, orthogonal.*

The notion of orthogonality and its relationship to design resolution clearly suggests that variance-optimal designs should also be available if the fitted response model contains first-order terms and one or more interaction terms. The following subsection addresses this topic.

### 8.3.2 Orthogonal Designs for Models Containing Interaction

As we indicated earlier, it is not uncommon for the process or system to require one or more interaction terms to accompany the linear main effect terms in a first-order model. The two-level factorials and fractions are quite appropriate. As in the case of the first-order model, the orthogonal design is variance-optimal—the sense again being that variances of all coefficients are minimized on a per observation basis. Based on the discussion earlier in Section 8.3, it should be apparent that the two-level full factorial is orthogonal for models containing main effects and any (or all) interactions involving cross products of linear terms. If a fraction is to be used, one must have sufficient resolution to ensure that no model terms are aliased. For example, if two factor interaction terms appear in the model, a resolution of at least V is appropriate.

As an illustration, suppose for  $k = 4$  the analyst postulates a model

$$y_i = \beta_0 + \sum_{i=1}^4 \beta_i x_i + \sum_{i < j=2}^4 \sum \beta_{ij} x_i x_j + \varepsilon$$

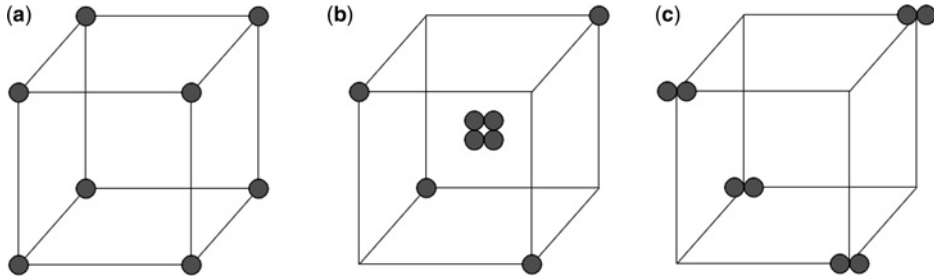
Then clearly no fractional factorial will be orthogonal, because a resolution IV fraction will result in aliasing two factor interactions with each other. In fact, a one-half fraction in this case would involve only eight design points, and the above model contains 11 model terms. A full  $2^4$  factorial is an orthogonal design for the above model.

Consider a second illustration in which five design variables are varied in a study in which the model can be written as

$$y = \beta_0 + \sum_{i=1}^5 \beta_i x_i + \sum_{i < j=2}^5 \sum \beta_{ij} x_i x_j + \varepsilon$$

In this case a complete  $2^5$  or a resolution V  $2^{5-1}$  serves as an orthogonal design. In this case, linear main effects are orthogonal to each other and orthogonal to two-factor interactions,





**Figure 8.2** Three designs for a first-order model in three design variables. (a)  $2^3$  factorial. (b)  $2^{3-1}$  fractional factorial with 4 center runs. (c) Replicated  $2^{3-1}$  fractional factorial.

**Effect of Center Runs** In Chapters 3 and 4 we discussed the utility of center runs as an important augmentation of the two-level factorial and fraction. This was discussed again and illustrated in Chapter 5. The addition of center runs to an orthogonal design does not alter the orthogonality property. However, the design is no longer a variance-optimal design; that is, the variances of regression coefficients are no longer minimized on a per observation basis. Obviously, the requirement that all levels be at  $\pm 1$  extremes is not met. Indeed, the center runs (runs at the zero coordinates for all factors in design units) add nothing to information about linear effects and interactions. However, let us not forget the importance of the detection of quadratic effects (see Chapter 3), a task that is handled very effectively and efficiently with the use of center runs.

**What Design Should Be Used?** The discussion of center runs and saturated designs in the context of orthogonal designs underscores a very important point. The design that is used in a given situation depends a great deal on the goal of the experiment and the assumptions that one is willing to make. It should also be clear that an optimal (i.e., variance-optimal) design is not always appropriate. Consider an example in which one is interested in fitting the model

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i \quad (i = 1, 2, 3, \dots, 8)$$

that is, a first-order model in three design variables. Among the candidate designs are three orthogonal designs that we will discuss and compare. All three of these designs involve eight experimental runs. Consider the following designs illustrated in Fig. 8.2:

**Design 1.** The  $2^3$  factorial

**Design 2.** Resolution III fraction of the  $2^3$  augmented with four center runs

**Design 3.** Resolution III fraction of the  $2^3$  with replicate runs at each design point

Designs 1 and 3 are both variance-optimal designs. The variances of regression coefficients will be  $\sigma^2/8$  for Designs 1 and 3. For Design 2, the variance of the intercept term will be  $\sigma^2/8$  but the variances of the linear coefficients will be  $\sigma^2/4$ . Comments concerning all three designs follow:

**Design 1.** It does not allow for estimation of replication error (pure error). There are four lack-of-fit degrees of freedom, but no error degrees of freedom are available

to test lack of fit. The design is appropriate if there is strong prior information that there is no model inadequacy from quadratic terms.

**Design 2.** Pure error degrees of freedom (3 df) are available for testing linear effects. One can also test for curvature (1 df for pure quadratic terms). The linear terms in the model are not estimated as precisely as for Designs 1 and 3. It should be used when one accepts the possibility of quadratic terms but can absolutely rule out interaction (the presence of pure quadratic terms with no interaction is not particularly likely).

**Design 3.** The design contains four pure error degrees of freedom but no lack-of-fit degrees of freedom associated with interaction terms or pure quadratic terms. Tests of significance can be made on model coefficients for the first-order model, but the design should not be used if either quadratic effects or interaction may be present. There is no model adequacy check of any kind.

From the above description it would seem that all these candidates possess disadvantages. However, one must keep in mind that with a sample size of eight and four model terms there is not much room for an efficient check for model inadequacy, because both lack-of-fit and pure error information is needed. Incidentally, a full  $2^3$  augmented with, say, three or more center runs would provide pure error degrees of freedom that could be used to test pure quadratic curvature and the presence of interaction.

### 8.3.3 Other First-Order Orthogonal Designs—The Simplex Design

The majority of applications of orthogonal designs for first-order models should suggest the use of the two-level designs that we have discussed here and in Chapters 3 and 4. However, it is important for the reader to be aware of a special class of first-order orthogonal designs that are saturated. This class is called the class of **simplex designs**. Let it be clear that these designs are saturated for a *first-order model*—a fact that implies that no information is available at all to study interactions or to examine lack of fit.

The main feature of the simplex design is that it requires  $N = k + 1$  observations to fit a first-order model in  $k$  variables. Geometrically, the design points represent the vertices of a regular sided figure. That is, the design points, given by the rows of the design matrix

$$\mathbf{D} = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & & \vdots \\ x_{1,N} & x_{2,N} & \cdots & x_{k,N} \end{bmatrix}$$

are points in  $k$  dimensions such that the angle that any two points make with the origin is  $\theta$ , where

$$\cos \theta = -\frac{1}{N-1} = -\frac{1}{k}$$

For  $k = 2$ ,  $N = 3$ ,  $\cos \theta = -\frac{1}{2}$ , and thus  $\theta = 120^\circ$ . Thus, for this case, the points are coordinates of an *equilateral triangle*. For  $k = 3$  and  $N = 4$ , the design points are the vertices of a

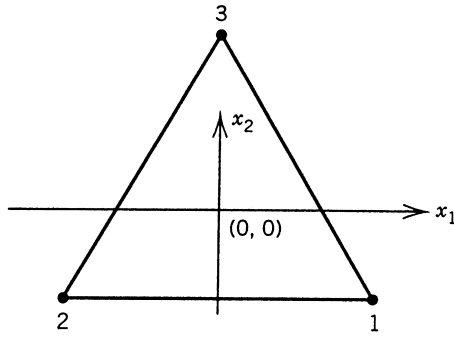


Figure 8.3 Design points for the simplex in two variables.

tetrahedron. For  $k = 2$ , the design matrix is given by

$$\mathbf{D} = \begin{bmatrix} x_1 & x_2 \\ \sqrt{3/2} & -1/\sqrt{2} \\ -\sqrt{3/2} & -1/\sqrt{2} \\ 0 & 2/\sqrt{2} \end{bmatrix} \tag{8.6}$$

Figure 8.3 reveals that the design for  $k = 2$  contains points on the vertices of an equilateral triangle. The three rows in the design matrix correspond to points 1, 2, and 3, respectively. The matrix  $\mathbf{X}$  is the design matrix augmented by a column of ones. Thus

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 & x_2 \\ 1 & \sqrt{3/2} & -1/\sqrt{2} \\ 1 & -\sqrt{3/2} & -1/\sqrt{2} \\ 1 & 0 & 2/\sqrt{2} \end{bmatrix} \tag{8.7}$$

As a result, one can readily observe the orthogonality from

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

For  $k = 3$ , the design involves the use of  $N = 4$  points and any two points make an angle  $\theta$  with the origin, where  $\cos \theta = -\frac{1}{3}$ . The appropriate design is given by

$$\mathbf{D} = \begin{bmatrix} x_1 & x_2 & x_3 \\ 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{8.8}$$



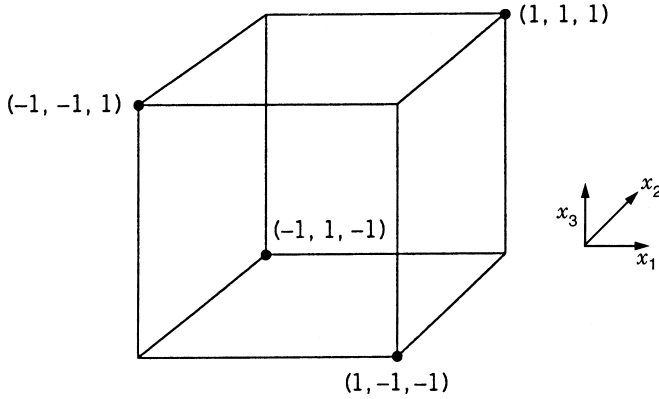


Figure 8.4 Simplex in three design variables.

If we use  $\mathbf{X} = [\mathbf{1D}]$ , it is clear that

$$\mathbf{X}'\mathbf{X} = 4\mathbf{I}_4$$

and hence the design is orthogonal for fitting the first-order model. Figure 8.4 shows the coordinates displayed in three dimensions. Note that any two points form the same angle with the origin  $(0, 0, 0)$ . Also note that in this special case the simplex is a one-half fraction of a  $2^3$ . The one we have chosen to display is the resolution III fraction with  $I = ABC$  as the defining relation. Indeed, the fraction formed from  $I = -ABC$  also is a simplex.

**Rotation and Construction of the Simplex** For a specific value of  $k$ , there are numerous simplex designs that can be constructed. In fact, if one rotates the rows of the design matrix through an angle, say  $\phi$ , the resulting design is still a simplex and thus still possesses the property of being first-order orthogonal. For example, a change in the orientation of the equilateral triangle of the design in Equation 8.6 will result in a second simplex. For  $k = 3$ , a design that is an alternative to that shown in Equation 8.8 is given by

$$\mathbf{D} = \begin{bmatrix} x_1 & x_2 & x_3 \\ 0 & \sqrt{2} & -1 \\ -\sqrt{2} & 0 & 1 \\ 0 & -\sqrt{2} & -1 \\ \sqrt{2} & 0 & 1 \end{bmatrix} \tag{8.9}$$

It is easily seen that this design is also first-order orthogonal, although from a pragmatic point of view the design in Equation 8.8 is usually preferable.

The construction of the matrix  $\mathbf{X}$  for the general  $k$ -dimensional simplex is quite simple. One merely starts with an orthogonal matrix  $\mathbf{O}$  of order  $N \times N$  with the elements of the first column being equal. The matrix  $\mathbf{X}$  is then

$$\mathbf{X} = \mathbf{O} \cdot N^{1/2}$$

Thus, it is easily seen that  $\mathbf{X}'\mathbf{X} = N(\mathbf{O}'\mathbf{O}) = N\mathbf{I}_N$ . For example, the  $k = 2$  simplex in Equation 8.7 can be constructed from the orthogonal matrix

$$\mathbf{O} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & 0 & 2 \end{bmatrix}$$

$$1/\sqrt{3} \quad 1/\sqrt{2} \quad 1/\sqrt{6}$$

The values at the bottom of each column are to be multiplied by each element in the column. We have

$$\mathbf{X} = \mathbf{O} \cdot \sqrt{3} = \begin{bmatrix} 1 & \sqrt{3/2} & -1/\sqrt{2} \\ 1 & -\sqrt{3/2} & -1/\sqrt{2} \\ 1 & 0 & 2/\sqrt{2} \end{bmatrix}$$

which is the matrix  $\mathbf{X}$  in Equation 8.7. The matrix  $\mathbf{X}$  given in the design of Equation 8.8 can be constructed from the simple orthogonal matrix

$$\mathbf{O} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$1/2 \quad 1/2 \quad 1/2 \quad 1/2$$

The reader should keep in mind that the simplex design is first-order saturated, which implies that there is no information available for lack of fit. If the analyst expects system interaction or pure quadratic curvature might be possible, the simplex is not an appropriate design.

**Example 8.3 Use of a Simplex Design** A simplex design is used in a laboratory experiment designed to build a first-order relationship between the growth ( $y$ ) of a particular organism and the percentage of glucose ( $x_1$ ), the percentage of yeast extract ( $x_2$ ), and the time in hours ( $x_3$ ) allowed for organism growth. In coded form the design variables are

$$x_1 = \frac{\text{glucose} - 30\%}{1.0}$$

$$x_2 = \frac{\text{yeast} - 0.5\%}{0.10}$$

$$x_3 = \frac{\text{time} - 45 \text{ hr}}{15}$$

Table 8.1 gives the design in terms of the original and coded variables and indicates the observed response.

The fitted equation is given by

$$\hat{y} = 9.54 + 0.97x_1 + 0.40x_2 + 1.59x_3$$

**TABLE 8.1 Experimental Data for Organism Growth Experiment Using Simplex Design**

| Run Number | Glucose $x_1$ (%) |             | Yeast $x_2$ (%) |             | Time, $x_3$ (hr) |       | Growth (g/liter) |
|------------|-------------------|-------------|-----------------|-------------|------------------|-------|------------------|
|            | Uncoded           | Coded       | Uncoded         | Coded       | Uncoded          | Coded |                  |
| 1          | 3.0               | 0           | 0.641           | $\sqrt{2}$  | 30               | -1    | 8.52             |
| 2          | 1.586             | $-\sqrt{2}$ | 0.500           | 0           | 60               | 1     | 9.76             |
| 3          | 3.0               | 0           | 0.359           | $-\sqrt{2}$ | 30               | -1    | 7.38             |
| 4          | 4.414             | $\sqrt{2}$  | 0.500           | 0           | 60               | 1     | 12.50            |

### 8.3.4 Definitive Screening Designs

Recall that a primary goal of a screening experiment often focuses on examining the contributions of different inputs to changes observed in the response. Hence good estimation of model parameters for a first-order model or a first-order model with interactions is an important criterion for evaluating the appropriateness of a design for this purpose.

Jones and Nachtsheim (2011a) introduce an important class of screening design that perform well for estimating first-order models with the potential to also estimate some of their two-way interactions. The designs also have the ability to do this while also balancing other objectives, such as allowing estimating of pure error and potential curvature from second-order terms. A key to how these designs manage to balance multiple design goals is to borrow from the ideas of multiple objective optimization described in Chapter 7. Recall, that one of the lessons from looking at multiple objectives is that often a small sacrifice on the performance of one goal (away from being optimal to near-optimal) can lead to disproportionate improvement on others.

Definitive screening designs (DSD) differ from other common screening designs in several important ways:

1. They use three levels of each of the factors—typically  $-1$ ,  $0$ , and  $+1$  in the scaled variables. This has the advantage over the traditional choice of two-level factors of allowing estimation of curvature from second-order terms, which previously had only been possible in a limited with the addition of center runs.
2. They use *nonregular* structures, as introduced in Section 4.8, that provide more flexible aliasing structures. In contrast, a fractional factorial has a *regular* structure where terms are either orthogonal or completely aliased. The nonregular structure means that terms in the model are not orthogonal, but rather have some level of correlation. The advantage of this approach is that it allows for more terms to be estimated, even if they cannot be completely independently assessed.

The resulting designs require only a moderate number of runs to explore different numbers of factors, and the analysis methods described in Section 4.9 for nonregular designs make interpretation of results straightforward.

Consider an experiment with four factors. The experimenter wishes to fit a model capable of examining all of the main effects and their two-factor interactions. Potentially, there may be some curvature from one or more of the factors, and hence it would be helpful to be able to estimate quadratic effects as well. Table 8.2 shows the levels of each factor for the design of size 11 of a DSD. Note how each factor has three levels, and there are four pairs

**TABLE 8.2 Definitive Screening Design for Four Factors**

| Run | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|-----|-------|-------|-------|-------|
| 1   | 0     | +1    | -1    | -1    |
| 2   | 0     | -1    | +1    | +1    |
| 3   | -1    | 0     | -1    | +1    |
| 4   | +1    | 0     | +1    | -1    |
| 5   | -1    | -1    | 0     | -1    |
| 6   | +1    | +1    | 0     | +1    |
| 7   | -1    | +1    | +1    | 0     |
| 8   | +1    | -1    | -1    | 0     |
| 9   | 0     | 0     | 0     | 0     |
| 10  | 0     | 0     | 0     | 0     |
| 11  | 0     | 0     | 0     | 0     |

of runs followed by three center runs. For each pair of runs (run 1–2, 3–4, etc.), one of the factors is set at 0, while the other factors are all at +1 or -1. The second run in the pair has the signs reversed from the first run for all nonzero entries for the remaining factors. In general, a DSD needs to have at least one center run, where additional ones help with improved estimation of pure error.

Recall that the experimenter wishes to estimate all of the main effects, but also consider their two-way interactions. For the four-factor case, this represents having a primary model of  $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon$ , but with interest in being able to examine the six two-factor interactions as well as the four quadratic terms.

To understand why this design is an advantageous one, consider the aliasing structure discussed in Section 8.2.2. Recall that if we assume a model of the form,  $y = \mathbf{X}_1 \beta_1 + \varepsilon$ , but the appropriate model turns out to be  $y = \mathbf{X}_1 \beta_1 + \mathbf{X}_2 \beta_2 + \varepsilon$ , the alias matrix  $A = (\mathbf{X}'_1 \mathbf{X}_1)^{-1} \mathbf{X}'_1 \mathbf{X}_2$  provides information about how the estimates  $b_1$  will be biased for  $\beta_1$ . In this case we define  $\mathbf{X}_1$  to be the  $\mathbf{X}$ -matrix corresponding to the four main effects, and  $\mathbf{X}_2$  to be the  $\mathbf{X}$ -matrix corresponding to the six two-factor interactions  $(x_1 x_2, x_1 x_3, \dots, x_3 x_4)$  and the four quadratic terms  $(x_1^2, \dots, x_4^2)$ . It is straightforward to show that  $A = \mathbf{0}$  for the DSD in Table 8.2. As a result, the main effects are independent of the two-factor interactions, and the main effects are independent of the quadratic effects. This means that the main effects can be estimated separately from any contributions from the other potential terms in the models, providing good understanding of the relative contributions of the factors.

While the DSDs were originally suggested as a screening design for primarily a first-order model, they are also powerful designs when some two-way interaction terms are also thought to be potentially active. Suppose that the experimenter for the four-factor experiment described in Table 8.2 wished to examine all of the main effects and their two-factor interactions. If all of these terms were thought to be active, then the DSD with 11 runs and 9 distinct design locations (since there are three center runs) would be too small to provide sufficient information to estimate all of the 11 terms in the model (one intercept, four main effects, and six two-factor interactions). However, for many experiments, there is an assumption of **effect sparsity**, as described in Section 3.5, which suggests that the experimenter believes that the system under study is dominated by a subset of the potential terms in the model, but does not know a priori which ones will be the dominant ones.

Using the nonregular design analysis methods described in Section 4.9, the experimenter can consider all of the main effects, two-factor interactions, and quadratic terms as candidates for the final model. Model selection techniques can identify the dominant terms, and using the pure error estimate from the center run replicates, some assessment of the statistical significance of these terms can be made. If some of the two-way interactions or quadratic terms are important to include in the model, then the correlation structure of these terms is moderate. See Jones and Nachtsheim (2011a) for more details.

DSDs are available for any number of factors with  $k \geq 4$ . For any  $k$ , the designs have the same structure with  $k$  pairs of runs plus one or more center runs. There are clearly a large number of different combinations of  $\pm 1$  that could be used for each of the nonzero factors in a given first run of the pair. Xiao, Lin, and Bai (2012) showed that a particular form of matrix, called a **conference matrix**, can be used to construct the set of  $k$  first runs for the design. To obtain the full design, the complement of each of the first runs is created (by reversing the sign of all nonzero entries), and one or more center runs are added. Hence the overall design would have the form

$$D = \begin{pmatrix} C \\ -C \\ 0 \end{pmatrix}$$

where  $C$  is the conference matrix,  $-C$  is its complement, and  $0$  is the set of one or more center runs.

The conference matrices for DSDs for  $k = 4, 5, 6, 7, 8$  are given below. For designs with  $k > 8$ , see Jones and Nachtsheim (2011a).

|   |  |   |
|---|--|---|
| <p><b>m = 4</b><br/>                 0+--<br/>                 -0-+<br/>                 --0-<br/>                 -++0</p>   | <p><b>m = 5</b><br/>                 0+---<br/>                 +0--+<br/>                 +-0+-<br/>                 +-+0+<br/>                 ++++0</p>   | <p><b>m = 6</b><br/>                 0+-----<br/>                 +0-++-<br/>                 --0+--<br/>                 -++0+-<br/>                 +-+-0-<br/>                 +++++-0</p> |
| <p><b>m = 7</b><br/>                 0+---+-<br/>                 -0+---+<br/>                 +-0++++<br/>                 +--0+--<br/>                 --++0--<br/>                 -+---+0+<br/>                 +++++-0</p> | <p><b>m = 8</b><br/>                 0-+++---<br/>                 -0-++++-<br/>                 --0+---+<br/>                 +-+0+---<br/>                 --+-0-+-<br/>                 +----+0++<br/>                 -++-++0+<br/>                 +++++-+0</p> |   |

The reader should verify that the construction method described for  $D$  above does correspond to Table 8.2 for  $k = 4$  with three center runs. Note that the four rows shown above correspond to runs 1, 3, 5 and 7.

**Odd Versus Even Numbers of Factors** One of the interesting features of the DSDs is the different numbers of factors have different characteristics. DSDs with an even number of factors have all of the main effect columns orthogonal to each other, allowing for

independent estimation and evaluation of their influence on the response. To see why this is not possible, consider the design matrix for the 5-factor DSD.

| Conference Matrix Order | Standardized Order |
|-------------------------|--------------------|
|                         | 0++++              |
|                         | +0---+             |
| C                       | +−0+−              |
|                         | +−+0+              |
|                         | <u>++++0</u>       |
|                         | 0---+              |
|                         | −0+++              |
| −C                      | −+0−+              |
|                         | −+−0−              |
|                         | <u>----0</u>       |
|                         | 00000              |

In order for any two main effects to be orthogonal to each other, there needs to be balance between the number of entries that lead to  $-1$  and  $+1$  when we take the product of the entries for the two columns. Consider whether columns 1 and 2 can be orthogonal. Looking at the standardized order of the DSD, the first four rows and the center run all contribute 0 to the sum. The remaining six rows contribute

$$(-1, -1, -1, -1, +1, +1)$$

Notably, each pair of runs contributes the same value, since the product of  $(+-)$  or  $(-+)$  lead to  $-1$  and products of  $(++)$  or  $(--)$  lead to  $+1$ . Hence for odd number of factor cases, we will have 4 + the number of center runs that contribute zero to the sum, and the remaining  $k - 2$  pairs (which must be odd) contribute an uneven number of  $-1$ s and  $+1$ s. Hence there will be a small amount of correlation between main effects in odd designs.

For even numbers of factors, there will again be  $k - 2$  pairs of runs that contribute nonzero values to the sum. In this case,  $k - 2$  is even (since  $k$  is even) with an overall design size for the non-center run rows that is a multiple of 4, and this allows for a balanced number of  $-1$ s and  $+1$ s to give an overall total of 0 to the sum of the product of the columns, leading to orthogonal effects.

Hence, DSDs with an even number of factors have slightly more desirable design properties; and if an experimenter finds that they naturally are interested in an odd number of factors, they might consider choosing a DSD with one more factor, to bump the design to be a multiple of 4. When the experimenter assigns the factors to the columns of the design, one of the columns can just be omitted for the missing factor.

Now consider a comparison between alternative designs for a first-order model with interactions for  $k = 4$ . Several choices might be considered suitable:

- (i) a  $2^4$  factorial ( $N = 16$ ),
- (ii) a  $2^{4-1}$  fractional factorial ( $N = 8$ ),
- (iii) a definitive screening design with 3 center runs ( $N = 8 + 3 = 11$ ),
- (iv) a  $2^4$  factorial with 3 center run ( $N = 16 + 3 = 19$ ),
- (v) a  $2^{4-1}$  fractional factorial with 3 center run ( $N = 11$ ).

We can compare these alternative designs on several criteria to see how they perform. Table 8.3 shows a summary of some advantages and disadvantages. As highlighted in Section 8.1, there are often noticeable trade-offs between competing alternatives. First, as is typically the case, if we increase the design size to the largest alternative, we are able to obtain a design able to estimate all of the terms in the first order with interactions model, with small variances for all of the estimated model coefficients,  $b_i$  and  $b_{ij}$ , and an estimate of natural variability from the center runs. However, if we are unwilling or unable to run such a large experiment, then we are forced to make some difficult trade-offs between alternatives. The DSD allows us to estimate all potential terms in the model, but at the expense of increase variances for the  $b_i$ s and  $b_{ij}$ s and nonorthogonality of the effects. However, these options may well be considered superior choices to the comparably sized fractional factorials, which have no option to de-alias pairs of two-factor interactions. The DSD is notably superior to all of the factorial and fractional factorial designs in its ability to provide separate estimates for curvature from the different factors. Notice how the addition of center runs provides an estimate of  $\sigma^2$ , but does not reduce the variance of the model parameters.

The comparison between alternative designs changes as the number of factors change, as different resolutions (IV and V) of designs become available with different aliasing structures. The size of the DSDs remains fixed at  $2k + n_c$ , while the factorial and fractional factorial designs fluctuates as different powers of 2 are available. For example, when  $k = 6$ , suitable options for designs include a Resolution VI  $2^{6-1}$  ( $N = 32 + n_c$ ), a Resolution IV  $2^{6-2}$  ( $N = 16 + n_c$ ), a Resolution III  $2^{6-3}$  with or without center runs ( $N = 8 + n_c$ ), all with or without center runs, and a DSD ( $N = 12 + n_c$ ). See Table 4.13 for more details on the defining equation for the various factorial designs. Note that while the  $2^{6-2}$  and  $2^{6-3}$  have problematic complete aliasing of key terms in the first-order with interactions model, the DSD is able to estimate all terms in the model, albeit with some nonorthogonality.

### 8.3.5 Another Variance Property—Prediction Variance

We earlier indicated that two-level orthogonal first-order designs with levels at  $\pm 1$  extremes result in variances of coefficients that are minimized on a per observation basis. This provided the inspiration for the term “variance optimal” that we used consistently as we discussed examples and applied the concept to models that involve interaction terms. Criteria involving variances of coefficients are certainly important. In fact, in Chapter 9 we explain the notion of special norms on  $(\mathbf{X}'\mathbf{X})^{-1}$  as we more formally deal with design optimality. However, we should also be attentive to the criterion of **prediction variance**. This concept was discussed in Chapter 2 in dealing with confidence intervals around  $\hat{y}(\mathbf{x})$  in regression problems, and again in Chapter 6 as a part of the total RSM analysis. However, the prediction variance concept is a very important aspect of the study of experimental design. It is important for the reader to become familiar with it in what is the simplest scenario—namely, the first-order model.

**Prediction Variance for the First-Order Model** Recall that the prediction variance, or variance of a predicted value, is given by

$$PV(\mathbf{x}) = \text{Var}[\hat{y}(\mathbf{x})] = \sigma^2 \mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)}$$

where  $\mathbf{x}^{(m)}$  is a function of the location in the design variables at which one predicts and also a function of the model. In fact, the  $(m)$  in  $\mathbf{x}^{(m)}$  reflects the **model**. In the case of a

TABLE 8.3 Comparison of Factorial, Fractional Factorial, and Definitive Screening Designs for Four Factors

| Design                 | Design Size, $N$ | First-Order $(X'X)^{-1}$   | First-Order + Interaction $(X'X)^{-1}$ | Curvature  | Aliasing of Interactions   | Estimation and Model Fitting of First-Order + Interactions                    |
|------------------------|------------------|----------------------------|--|--|--|---|
| $2^4$                  | 16               | $\frac{1}{16}\mathbf{I}_5$ | $\frac{1}{16}\mathbf{I}_{11}$          | Not possible   | None—all effects orthogonal  | Half-normal plot—all terms estimable  |
| $2^{4-1}$              | 8                | $\frac{1}{8}\mathbf{I}_5$  | Not possible                           | Not possible   | All two-way interactions completely aliased with another effect                              | Half-normal plot—unable to separate pairs of interaction effects              |
| $\text{DSD} + n_c = 3$ | 11               | $\frac{1}{6}\mathbf{I}_5$  | Not possible                           | Can estimate each term separately                    | Main effects orthogonal to each other & two-way interactions, two-way interactions estimable | Forward stepwise regression with estimate of natural variation                |
| $2^4 + n_c = 3$        | 19               | $\frac{1}{16}\mathbf{I}_5$ | $\frac{1}{16}\mathbf{I}_{11}$          | Can estimate combined quadratic for all four factors | None—all effects orthogonal  | Analysis of variance (ANOVA) with estimate of natural variation               |
| $2^{4-1} + n_c = 3$    | 11               | $\frac{1}{8}\mathbf{I}_5$  | Not possible                           | Can estimate combined quadratic for all four factors | Same as $2^{4-1}$  | ANOVA with estimate of natural variation—unable to separate interaction pairs |



strictly first-order model, we obtain

$$\mathbf{x}^{(1)'} = [1, x_1, x_2, \dots, x_k]$$

For a  $k = 2$  model containing  $x_1, x_2$  and their interaction, we obtain

$$\mathbf{x}^{(1)'} = [1, x_1, x_2, x_1x_2]$$

One can easily see that  $\text{Var}[\hat{y}(\mathbf{x})]$  varies from location to location in the design space. In addition, the presence of  $(\mathbf{X}'\mathbf{X})^{-1}$  attests to the fact that the criterion is very much a function of the experimental design. The reader should view the prediction variance as a reflection of how well one predicts with the model, given a particular choice of design.

In studies that are done to compare designs, it is often convenient to **scale** the prediction variance, that is, work with the **scaled prediction variance**

$$\text{SPV}(\mathbf{x}) = \frac{N \text{Var}[\hat{y}(\mathbf{x})]}{\sigma^2} = N\mathbf{x}^{(m)'}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}^{(m)} \tag{8.10}$$

The division by  $\sigma^2$  makes the quantity scale-free, and the multiplication by  $N$  allows the quantity to reflect variance on a *per observation basis*. That is, if two designs are being compared, the scaling by  $N$  automatically “punishes” the design with the larger sample size. It forces a premium on efficiency.

The scaled prediction variance in Equation 8.10 is quite simple for the case of a variance optimal first-order orthogonal design. In fact, because  $(\mathbf{X}'\mathbf{X})^{-1} = (1/N)\mathbf{I}_p$ , we have

$$\begin{aligned} \text{SPV}(\mathbf{x}) &= [1, x_1, x_2, \dots, x_k] \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = 1 + \sum_{i=1}^k x_i^2 \\ &= 1 + \rho_{\mathbf{x}}^2 \end{aligned} \tag{8.11}$$

where  $\rho_{\mathbf{x}}$  is the distance that the point  $\mathbf{x}' = [x_1, x_2, \dots, x_k]$  is away from the design origin. As a result, the SPV on  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2 = 1$  at the design center and becomes larger as one moves toward the design perimeter. Anyone who recalls confidence intervals on  $E[y(\mathbf{x})]$  in linear regression should not be surprised at this result. As an illustration, for a  $2^3$  factorial design, the following are values of  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  at different locations where one may wish to predict:

| Location                                | $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ |
|---|--|
| 0, 0, 0                                 | 1  |
| $\frac{1}{2}, \frac{1}{2}, \frac{1}{2}$ | 1.75   |
| $\frac{1}{2}, \frac{1}{2}, 1$           | 2.50   |
| $\frac{1}{2}, 1, 1$                     | 3.25   |
| 1, 1, 1                                 | 4  |

Thus, the scaled prediction variance increases *fourfold* as one moves from the design center to the design perimeter. For, say, a  $2^5$  factorial at the perimeter [i.e.,  $(\pm 1, \pm 1, \pm 1, +1, \pm 1)$ ],  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2 = 6$  and hence the variance increases *sixfold*.

In the case of a lesser design (i.e., one that is not variance-optimal, or even orthogonal), this change in prediction variance is even more pronounced. For example, consider a one-half fraction (resolution III) of a  $2^3$  factorial with four center runs as an alternative to a  $2^3$ . We know that this design is orthogonal but not variance-optimal. Recall this design as Design 2 in the comparisons we made in Section 8.3.2. The  $2^3$  was Design 1. For Design 2 the model matrix is

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{X}'\mathbf{X} = \begin{bmatrix} 8 & 0 & 0 & 0 \\ & 4 & 0 & 0 \\ & & 4 & 0 \\ & & & 4 \end{bmatrix}$$

with the variance-covariance matrix (apart from  $\sigma^2$ ) being

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} \frac{1}{8} & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{bmatrix}$$

Consequently,

$$\begin{aligned} \text{SPV}(\mathbf{x}) &= 8[1, x_1, x_2, x_3] \begin{bmatrix} \frac{1}{8} & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= 1 + 2(x_1^2 + x_2^2 + x_3^2) \\ &= 1 + 2\rho_x^2 \end{aligned}$$

As a result, it is apparent that  $\text{SPV}(\mathbf{x})$  for Design 2 increases much faster as one approaches the design perimeter than does Design 1. In fact, at  $(\pm 1, \pm 1, \pm 1)$ ,  $\text{SPV}(\mathbf{x}) = 4$  for Design 1 and  $\text{SPV}(\mathbf{x}) = 7$  for Design 2. Incidentally, Design 3, which is a resolution III fraction of a  $2^3$  with replicate runs, has prediction variance properties identical to that of Design 1; they are both variance-optimal designs.

The reader should view prediction variance as another variance-type criterion by which comparisons among designs can be made, where here the focus is on good prediction of new observations. It is different than merely comparing variances of individual coefficients.

Much more attention will be given to  $SPV(\mathbf{x}) = N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  later in this chapter and in future chapters.

## 8.4 DESIGNS FOR FITTING SECOND-ORDER MODELS

Before we embark on the large topic of designed experiments for second-order models, we should review for the reader some minimum design requirements and the RSM philosophy that motivates one or two of the classes of designs that we will subsequently present in detail. Variable screening is an essential phase of RSM that was presented at length in Chapter 1. Here, of course, the two-level factorial designs and fractions play a major role. Any sequential movement (region seeking, such as the Path of Steepest Ascent described in Chapter 5) that is necessary is also accomplished with a first-order design. Of course, there may be instances where region-seeking via steepest ascent (or descent) and/or variable screening are not required. However, the possibility of either or both should be included in the total sequential plan. At some point the researcher will be interested in fitting a second-order response surface in the design variables  $x_1, x_2, \dots, x_k$ . This response surface analysis may involve optimization as discussed in Chapter 6. It may lead to even more sequential movement through the use of canonical or ridge analysis. But, regardless of the form of the analysis, the purpose of the experimental design is one that should allow the user to fit the second-order model

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j=2}^k \sum_{j=2}^k \beta_{ij} x_i x_j + \varepsilon \quad (8.12)$$

The model of Equation 8.12 contains  $1 + 2k + k(k-1)/2$  parameters. There must be at least this number of distinct design points and at least three levels of each design variable. Now, of course, these represent minimum conditions, and one should keep in mind the 11 desirable design characteristics listed at the beginning of this chapter. In what follows in this chapter and in Chapter 9 we discuss important design properties for the second-order model and specific classes of second-order designs. The reader will recall from the previous section that in the case of first-order designs (or first-order-with-interaction designs), a dominant property is orthogonality. In the case of second-order designs, orthogonality ceases to be such an important issue, and estimation of individual coefficients, while still important, becomes secondary to the scaled prediction variance,  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ . This stems from the fact that there is often less concern with what variables belong in the model than with the quality of  $\hat{y}(\mathbf{x})$  as a prediction or, rather, an estimator for  $E[y(\mathbf{x})]$ . We now introduce, formally, the class of central composite designs.

### 8.4.1 The Class of Central Composite Designs

The **central composite designs** (CCDs) were introduced in an informal way in Chapter 6. Some of the examples used for illustrative purposes involved the use of the CCD. The CCD is without a doubt the most popular class of second-order designs. It was introduced by Box and Wilson (1951).

Much of the motivation of the CCD evolves from its use in **sequential experimentation**. It involves the use of a **two-level factorial** or fraction (resolution V) combined with the following  $2k$  **axial** or **star** points:

| $x_1$     | $x_2$     | ... | $x_k$     |
|-----------|-----------|-----|-----------|
| $-\alpha$ | 0         | ... | 0         |
| $\alpha$  | 0         | ... | 0         |
| 0         | $-\alpha$ | ... | 0         |
| 0         | $\alpha$  | ... | 0         |
| $\vdots$  | $\vdots$  |     | $\vdots$  |
| 0         | 0         | ... | $-\alpha$ |
| 0         | 0         | ... | $\alpha$  |

As a result, the design involves, say,  $F$  factorial points,  $2k$  axial points, and  $n_c$  center runs. The sequential nature of the design becomes very obvious. The factorial points represent a variance-optimal design for a first-order model or a first-order + two-factor interaction model. Center runs clearly provide information about the existence of curvature in the system. If curvature is found in the system, the addition of axial points allow for efficient estimation of the pure quadratic terms.

While the genesis of this design is derived from sequential experimentation, the CCD is a very efficient design in situations that call for a nonsequential batch response surface experiment. In effect, the three components of the design play important and somewhat different roles.

1. The resolution V fraction contributes substantially to the estimation of linear terms and two-factor interactions. It is variance-optimal for these terms. The factorial points are the only points that contribute to the estimation of the interaction terms.
2. The axial points contribute in a large way to estimation of quadratic terms. Without the axial points, only the sum of the quadratic terms,  $\sum_{i=1}^k \beta_{ii}$ , can be estimated. The axial points do not contribute to the estimation of interaction terms.
3. The center runs provide an internal estimate of error (pure error) and contribute toward the estimation of quadratic terms.

The areas of flexibility in the use of the central composite design reside in the selection of  $\alpha$ , the axial distance, and  $n_c$ , the number of center runs. The choice of these two parameters can be very important. The choice of  $\alpha$  depends to a great extent on the region of operability and region of interest. The choice of  $n_c$  often has an influence on the distribution of  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  in the region of interest. More will be said about the choice of  $\alpha$  and  $n_c$  in subsequent sections. Figures 8.5 and 8.6 show the CCD for  $k=2$  and  $k=3$ . For the  $k=2$  case the value of  $\alpha$ , the axial distance is  $\sqrt{2}$ . For  $k=3$ , the value of  $\alpha$  is  $\sqrt{3}$ . Note that for  $k=3$  the axial points come through the six faces at a distance  $\sqrt{3}$  from the origin. For  $k=2$  the design represents eight points equally spaced on a circle, plus the center runs. For  $k=3$  the design represents 14 points all on a common sphere, plus center runs.

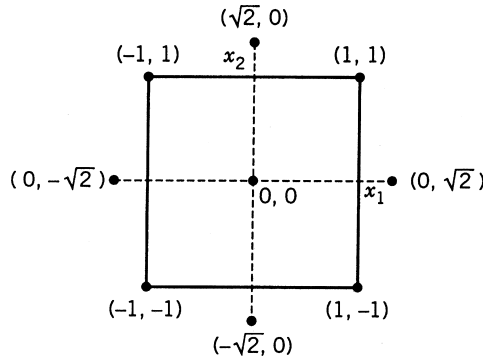


Figure 8.5 Central composite design for  $k = 2$  and  $\alpha = \sqrt{2}$ .

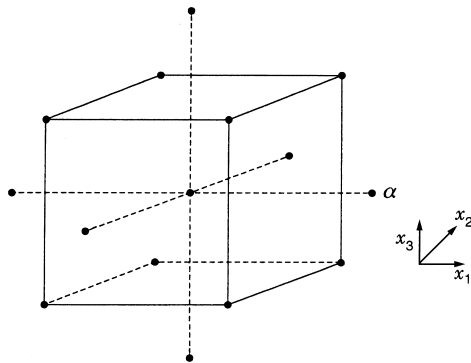


Figure 8.6 Central composite design for  $k = 3$  and  $\alpha = \sqrt{3}$ .

The value of the axial distance generally varies from 1.0 to  $\sqrt{k}$ , the former placing all axial points on the face of the cube or hypercube, the latter resulting in all points being placed on a common sphere. There are times when two or more center runs are needed and times when one or two will suffice. We will allocate considerable space in Chapter 9 to comparison of the CCD with other types of designs. We will discuss the choice of  $\alpha$  and  $n_c$  following a numerical example of the use of a CCD.

In the example that follows we use a real-life numerical example to gain insight into the structure of the  $\mathbf{X}$  and  $\mathbf{X}'\mathbf{X}$  matrices for the general case of a CCD.

**Example 8.4 The Breadwrapper Experiment** An experiment was conducted to study the response surface relating the strength of breadwrapper stock in grams per square inch to sealing temperature ( $x_1$ ), cooling bar temperature ( $x_2$ ), and percent polyethylene additive ( $x_3$ ). The definition of the design levels are [see Myers (1976)]:

$$x_1 = \frac{\text{temp.} - 255^\circ\text{F}}{30}$$

$$x_2 = \frac{\text{temp.} - 55^\circ\text{F}}{9}$$

$$x_3 = \frac{\text{polyethylene} - 1.1\%}{0.6}$$

Five levels of each factor are involved in this design. The coded and natural levels are given by the following:

|       | -1.682 | -1.000 | 0.000 | 1.000 | 1.682 |
|-------|--------|--------|-------|-------|-------|
| $x_1$ | 204.5  | 225    | 255   | 285   | 305.5 |
| $x_2$ | 39.9   | 46     | 55    | 64    | 70.1  |
| $x_3$ | 0.09   | 0.5    | 1.1   | 1.7   | 2.11  |

The design matrix **D** and the vector **y** of responses are

$$\mathbf{D} = \begin{matrix} & x_1 & x_2 & x_3 \\ \begin{bmatrix} -1 & -1 & -1 \\ 1 & -1 & -1 \\ -1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \\ -1.682 & 0 & 0 \\ 1.682 & 0 & 0 \\ 0 & -1.682 & 0 \\ 0 & 1.682 & 0 \\ 0 & 0 & -1.682 \\ 0 & 0 & 1.682 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & , & \mathbf{y} = \begin{bmatrix} 6.6 \\ 6.9 \\ 7.9 \\ 6.1 \\ 9.2 \\ 6.8 \\ 10.4 \\ 7.3 \\ 9.8 \\ 5.0 \\ 6.9 \\ 6.3 \\ 4.0 \\ 8.6 \\ 10.1 \\ 9.9 \\ 12.2 \\ 9.7 \\ 9.7 \\ 9.6 \end{bmatrix}
 \end{matrix}$$

Note that the  $k = 3$  CCD for this example uses  $\alpha = 1.682$  with  $n = 6$  center runs. While this is not the axial distance that stretches to the faces of the cube it is a value to which some importance is attached. It allows the design to possess the property of rotatability (defined in Section 8.2) and thus the derivation of the value  $\alpha = 1.682$  for the  $k = 3$  breadwrapper design will be discussed in Section 8.4.2. We shall use the breadwrapper design to highlight the structure of **X** and **X'X**. For the design matrix **D** and the second-order model:

|   | $x_1$ | $x_2$    | $x_3$ | $x_1^2$ | $x_2^2$ | $x_3^2$ | $x_1x_2$ | $x_1x_3$ | $x_2x_3$ |
|---|-------|----------|-------|---------|---------|---------|----------|----------|----------|
| 1 |       |          |       | 1       | 1       | 1       | 1        | 1        | 1        |
| 1 |       |          |       | 1       | 1       | 1       | -1       | -1       | 1        |
| 1 |       |          |       | 1       | 1       | 1       | -1       | 1        | -1       |
| 1 |       |          |       | 1       | 1       | 1       | 1        | -1       | -1       |
| 1 |       |          |       | 1       | 1       | 1       | 1        | -1       | -1       |
| 1 |       |          |       | 1       | 1       | 1       | -1       | 1        | -1       |
| 1 |       |          |       | 1       | 1       | 1       | -1       | -1       | 1        |
| 1 |       |          |       | 1       | 1       | 1       | 1        | 1        | 1        |
| 1 |       | <b>D</b> |       | 2.828   | 0       | 0       | 0        | 0        | 0        |
| 1 |       |          |       | 2.828   | 0       | 0       | 0        | 0        | 0        |
| 1 |       |          |       | 0       | 2.828   | 0       | 0        | 0        | 0        |
| 1 |       |          |       | 0       | 2.828   | 0       | 0        | 0        | 0        |
| 1 |       |          |       | 0       | 0       | 2.828   | 0        | 0        | 0        |
| 1 |       |          |       | 0       | 0       | 2.828   | 0        | 0        | 0        |
| 1 |       |          |       | 0       | 0       | 0       | 0        | 0        | 0        |
| 1 |       |          |       | 0       | 0       | 0       | 0        | 0        | 0        |
| 1 |       |          |       | 0       | 0       | 0       | 0        | 0        | 0        |
| 1 |       |          |       | 0       | 0       | 0       | 0        | 0        | 0        |
| 1 |       |          |       | 0       | 0       | 0       | 0        | 0        | 0        |
| 1 |       |          |       | 0       | 0       | 0       | 0        | 0        | 0        |
| 1 |       |          |       | 0       | 0       | 0       | 0        | 0        | 0        |

(8.13a)

| $b_0$ | $b_1$     | $b_2$  | $b_3$  | $b_{11}$ | $b_{22}$ | $b_{33}$ | $b_{12}$ | $b_{13}$ | $b_{23}$ |
|-------|-----------|--------|--------|----------|----------|----------|----------|----------|----------|
| 20    | 0         | 0      | 0      | 13.658   | 13.658   | 13.658   | 0        | 0        | 0        |
|       | 13.658    | 0      | 0      | 0        | 0        | 0        | 0        | 0        | 0        |
|       |           | 13.658 | 0      | 0        | 0        | 0        | 0        | 0        | 0        |
|       |           |        | 13.658 | 0        | 0        | 0        | 0        | 0        | 0        |
|       |           |        |        | 24       | 8        | 8        | 0        | 0        | 0        |
|       |           |        |        |          | 24       | 8        | 0        | 0        | 0        |
|       |           |        |        |          |          | 24       | 0        | 0        | 0        |
|       |           |        |        |          |          |          | 8        | 0        | 0        |
|       |           |        |        |          |          |          |          | 8        | 0        |
|       |           |        |        |          |          |          |          |          | 8        |
|       | Symmetric |        |        |          |          |          |          |          |          |

(8.13b)

Note that in the  $\mathbf{X}$  matrix the first four columns involving the column of ones and the linear term columns are mutually orthogonal as one would expect since they jointly represent columns of an orthogonal design. The same is true with the last three columns containing the columns that reflect presence of interaction terms. In addition, these two sets of columns are mutually orthogonal to one another. This is a result of the presence of the factorial portion of the CCD and the fact that that axial are chosen such that orthogonality is maintained among the linear columns. This produces a large number of zeros on the off diagonal elements of  $\mathbf{X}'\mathbf{X}$ . In fact the only non-zero entries on the off diagonals, the values of 8, exist because the “pure” quadratic terms are, of course, not mutually orthogonal. [The nonzeros in the first row and column (i.e., 13.658) could be made zero by centering the  $\mathbf{x}_i^2$  values

**TABLE 8.4 Analysis of Variance for Breadwrapper Stock Data**

| Source of Variation               | Sum of Squares | Degrees of Freedom | Mean Square | <i>F</i> |
|-----------------------------------|----------------|--------------------|-------------|----------|
| Regression (linear and quadratic) | 70.3056        | 9                  | 7.8117      | 7.87     |
| Lack of fit                       | 6.9044         | 5                  | 1.3809      | 1.39     |
| Error                             | 4.9600         | 5                  | 0.9920      |          |
| Total                             | 82.1700        | 19                 |             |          |

in the model.] The presence of many zeros on the off diagonals contribute to making the CCD a very efficient design. In addition, we will observe in later sections that the value of  $n_c$ , the number of center runs can be a very important design parameter.

The  $\mathbf{X}'\mathbf{X}$  matrix in Equation 8.13b is certainly specific to the conditions of the breadwrapper problem. However, it is important to understand that the zeros appear in the corresponding positions for any CCD for the reasons described above. In sections that follow we will revisit the general form of  $\mathbf{X}'\mathbf{X}$  as discussed for the CCD here with regard to how it effects the variance of prediction, that is, specifically  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ . Indeed certain values of the diagonals and the off-diagonal elements can dramatically influence the distribution of this scaled prediction variance quantity inside the design region of the CCD.

The least squares procedure gives the estimated second-order response function:

$$\hat{y} = 10.165 - 1.1036x_1 + 0.0872x_2 + 1.020x_3 - 0.760x_1^2 - 1.042x_2^2 - 1.148x_3^2 - 0.350x_1x_2 - 0.500x_1x_3 + 0.150x_2x_3$$

An analysis of variance is shown in Table 8.4. Note that there are five degrees of freedom for lack of fit, representing contributions from third-order terms. The *F*-test for lack of fit is not significant.

The stationary point is computed based on the methodology discussed in Chapter 6:

$$\begin{aligned} \mathbf{x}_s &= -\frac{1}{2}\hat{\mathbf{B}}^{-1}\mathbf{b} \\ &= \begin{bmatrix} -1.011 \\ 0.260 \\ 0.681 \end{bmatrix} \end{aligned}$$

with the predicted response at the stationary point given by  $\hat{y}(\mathbf{x}_s) = 11.08$ .

The eigenvalues of the matrix  $\hat{\mathbf{B}}$  are found to be

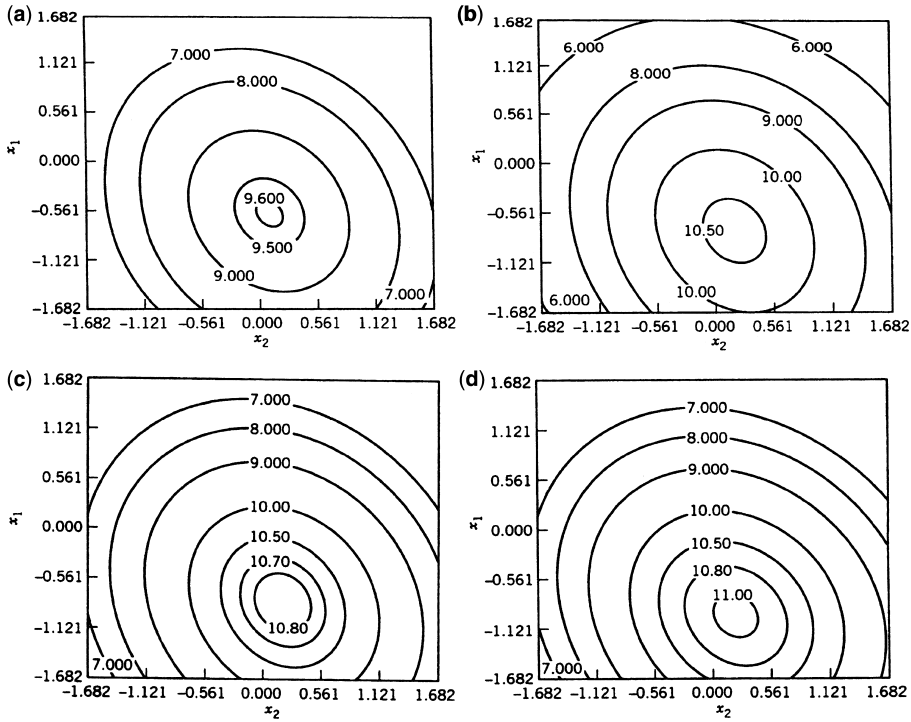
$$\lambda_1 = -0.562, \quad \lambda_2 = -1.271, \quad \lambda_3 = -1.117$$

Thus the canonical form is given by

$$\hat{y} = 11.08 - 0.562w_1^2 - 1.271w_2^2 - 1.117w_3^2$$

As a result, the stationary point is a point of estimated maximum mean strength of the breadwrapper. Figure 8.7 displays contour graphs that show the flexibility available around the estimated optimum. Contours of constant response are shown for  $x_3$  (percent polyethylene) values of  $-0.5, 0.00, 0.25, \text{ and } 0.5$ . The purpose is to determine how much strength is





**Figure 8.7** Contours of constant strength at different levels of the percentage of polyethylene,  $x_3$ . (a)  $x_3 = -0.5$ . (b)  $x_3 = 0$ . (c)  $x_3 = 0.25$ . (d)  $x_3 = 0.5$ .

lost by moving percent polyethylene  $x_3$  off the optimum value of 0.681 (coded). From the contour plots, it becomes apparent that even if the polyethylene content is reduced to a value as low as 0.25 (coded), the estimated maximum strength is reduced to only slightly less than 10.9 psi. A reduction to 0.500 in polyethylene will still produce a maximum that exceeds 11.0 psi.

The objective of this section is to formally introduce the user to the CCD and to allow more insight into its properties. However, the CCD is such an important part of the heritage and practical use of RSM that it will be revisited frequently in the text. In the next section we reintroduce the notion of **prediction variance**—that is,  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  in the case of second-order models. Simultaneously, we discuss the notion of the property of **rotatability**. This will allow the discussion of the choice of  $\alpha$  and  $n_c$  to be connected to design properties.

### 8.4.2 Design Moments and Property of Rotatability

Many of the properties of experimental designs are connected to the manner in which the points are *distributed in the region* of experimentation. Specifically, this distribution of points in space has a profound effect on the distribution of  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ , the scaled prediction variance. The distribution of design points is nicely quantified by its **design moments**. The term moments has the same conceptual meaning as the term sample moments that is taught in elementary statistics. We learn early in our training that the nature of the sample of the data is well characterized by its moments—for example, sample mean (first

moment) and sample variance (second moment). We also recall that symmetry in a sample is quantified by the third moment. In the case of RSM, the moments that reflect important geometry in the design must also be a function of the model being fit.

Indeed the important moments come from the **moment matrix**

$$\mathbf{M} = \frac{\mathbf{X}'\mathbf{X}}{N}$$

For example, a  $2^k$  factorial or fractional factorial design for a first-order model is orthogonal and its moment matrix is easily seen to be

$$\mathbf{M} = \frac{\mathbf{X}'\mathbf{X}}{N} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \end{bmatrix} = \mathbf{I}_k$$

The off-diagonal moments are zero due to the orthogonality of the columns of the  $\mathbf{X}$  matrix, and the diagonal moments are all equal to 1, since the  $\mathbf{X}$  matrix is given by

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1N} & x_{2N} & \cdots & x_{kN} \end{bmatrix}$$

where each of the  $x_{ij}$  entries are  $\pm 1$ . We then define the moments as *first moments*; that is,  $\frac{1}{N} \sum_{u=1}^N x_{iu}$ , the *second mixed moments*  $\frac{1}{N} \sum_{u=1}^N x_{iu}x_{ju}$  for  $i \neq j$  and the *second pure moments*  $\frac{1}{N} \sum_{u=1}^N x_{iu}^2$ . The first and second mixed moments are called **odd moments**, since they have at least one variable with an odd power. These are clearly zero for this design. The  $k$  second pure moments are then **even moments** and, of course, are equal to 1 in this case. It should be clear that the second mixed moments are analogous to a sample covariance for the sample moments case in basic statistics, where the first moment can be viewed like a sample mean, and the second pure moments as sample variances.

The design moments that carry importance in characterizing variance properties of a design are clearly a function of the order of the model. In the case of the  $2^k$  factorial or fractional factorial design illustrated above, the moments through order 2 are important since  $\mathbf{M} = \mathbf{X}'\mathbf{X}/N$  contains moments through order 2. However, in the case of the CCD with the corresponding second-order design,  $\mathbf{M}$  contains moments through order 4. This is evident when we observe  $\mathbf{X}$  and  $\mathbf{X}'\mathbf{X}$  in Equations 8.13a and 8.13b, respectively. In fact from (8.13b), the relevant moments are

$$\begin{aligned} [i] &= \frac{1}{N} \sum_{u=1}^N x_{iu} && \text{first moments} \\ [ii] &= \frac{1}{N} \sum_{u=1}^N x_{iu}^2 && \text{second pure moments} \\ [ij] &= \frac{1}{N} \sum_{u=1}^N x_{iu}x_{ju} && \text{second mixed moments} \end{aligned}$$

$$\begin{aligned}
 [iii] &= \frac{1}{N} \sum_{u=1}^N x_{iu}^3 && \text{third pure moments} \\
 [ijj] &= \frac{1}{N} \sum_{u=1}^N x_{iu}^2 x_{ju}, \quad [ijk] = \frac{1}{N} \sum_{u=1}^N x_{iu} x_{ju} x_{ku} && \text{third mixed moments} \\
 [iiii] &= \frac{1}{N} \sum_{u=1}^N x_{iu}^4 && \text{fourth pure moments} \\
 [iiij] &= \frac{1}{N} \sum_{u=1}^N x_{iu}^3 x_{ju}, \quad [iijj] = \frac{1}{N} \sum_{u=1}^N x_{iu}^2 x_{ju}^2 && \text{fourth mixed moments} \\
 [iijk] &= \frac{1}{N} \sum_{u=1}^N x_{iu}^2 x_{ju} x_{ku}, \quad [ijkl] = \frac{1}{N} \sum_{u=1}^N x_{iu} x_{ju} x_{ku} x_{lu}
 \end{aligned}$$

It is clear that for the CCD all odd moments through order four—that is, moments that contain at least one odd power, that is,  $[i]$ ,  $[ij]$ ,  $[iii]$ ,  $[iiij]$ ,  $[ijk]$ ,  $[iijj]$ , and  $[iijk]$ —are zero for  $i \neq j \neq k$ . Values of zero for odd moments suggest a certain design symmetry which is certainly apparent when one observes the design in Figs. 8.5 and 8.6. In fact from the  $\mathbf{X}'\mathbf{X}$  matrix in Equation 7.13b, the only nonzero moments for the  $k = 3$  CCD are  $[ii]$ ,  $[iijj]$ , and  $[iiii]$  for all  $i \neq j$ .

The important variance properties of an experimental design are determined by the nature of the moment matrix. This should be evident because the matrix  $(\mathbf{X}'\mathbf{X})^{-1}$  and hence  $\mathbf{X}'\mathbf{X}$  are so important in characterizing variance properties—that is, variances and covariances of regression coefficients as well as prediction variance.

The importance of design moments and hence the moment matrix in characterizing the scaled prediction variance, SPV, is easily seen by observing the definition of SPV in Equation 8.10.

$$\begin{aligned}
 \text{SPV}(\mathbf{x}) &= N \mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)} \\
 &= \mathbf{x}^{(m)'} \left( \frac{\mathbf{X}'\mathbf{X}}{N} \right)^{-1} \mathbf{x}^{(m)} \\
 &= \mathbf{x}^{(m)'} \mathbf{M}^{-1} \mathbf{x}^{(m)}
 \end{aligned}$$

So the SPV is a quadratic form of the moment matrix,  $\mathbf{M}$ . Thus the nature of the design moments has a profound effect on the SPV and its distribution in the design space.

It is important for a second-order design to possess a reasonably stable distribution of the scaled prediction variance  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  throughout the experimental design region. It must be clearly understood that the experimenter does not know at the outset where in the design space he or she may wish to predict, or where in the design space the optimum may lie. Thus, a reasonably stable scaled prediction variance  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  provides insurance that the quality of the  $\hat{y}(\mathbf{x})$  as a prediction of future response values is similar throughout the region of interest. To this end, Box and Hunter (1957) developed the notion of **design rotatability**.

A **rotatable** design is one for which  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  has the same value at any two locations that are the same distance from the design center. In other words,  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  is constant on spheres.

The purpose of design rotatability was, in part, to impose a type of stability on  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ . The rationale of rotatability is that at two locations in the design space  $\mathbf{x}_1$  and  $\mathbf{x}_2$  for which the distances from the origin are the same [i.e.,  $(\mathbf{x}'_1\mathbf{x}_1)^{1/2} = (\mathbf{x}'_2\mathbf{x}_2)^{1/2}$ ], the predicted values  $\hat{y}(\mathbf{x}_1)$  and  $\hat{y}(\mathbf{x}_2)$  should be equally good—that is, have equal variance. While rotatability itself does not ensure stability or even near-stability throughout the design region, in many cases it provides some useful guidelines for the choice of design parameters—for example, the choice of  $\alpha$  and  $n_c$  in the CCD. The importance of rotatability as a design property depends on many things, not the least of which is the nature of the region of interest and region of operability. It is important to note that rotatability or near-rotatability is often easy to achieve without the sacrifice of other important design properties. In what follows, we present the foundation that will allow the determination of necessary and sufficient conditions for design rotatability.

**Moment Matrix for a Rotatable Design (First and Second Order)** In this section we give the moment matrix for a rotatable design for both first- and second-order models. Additional theoretical development regarding the general case is given in Appendix 1. For the case of a **first-order model**, a design is rotatable if and only if *odd moments through order two are zero and the pure second moments are all equal*. In other words,

$$\begin{aligned} [i] &= 0 & (i = 1, 2, \dots, k) \\ [ij] &= 0 & (i \neq j, i, j = 1, 2, \dots, k) \\ [ii] &= \lambda_2 & (i = 1, 2, \dots, k) \end{aligned}$$

The quantity  $\lambda_2$  is determined by the scaling of the design. As a result in the first-order case, the moment conditions for a rotatable design are equivalent to the moment conditions for a variance-optimal design. Indeed, with scaling that allows levels at  $\pm 1$  in a two-level design of resolution III or higher, we have

$$[i] = 0, \quad [ij] = 0, \quad [ii] = 1.0 \quad (i = 1, 2, \dots, k, i \neq j)$$

In other words,  $\lambda_2$  is set at 1.0 due to the standard  $\pm 1$  scaling. The equivalence of rotatability to variance optimality in the first-order case should not be surprising. The reader should recall the result of Equation 8.11. We showed that

$$\text{SPV}(\mathbf{x}) = 1 + \rho_{\mathbf{x}}^2$$

for a first-order variance optimal design. Thus, the scaled prediction variance  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  is a function of  $\mathbf{x}$  only through  $\rho_{\mathbf{x}}$ , the distance of  $\mathbf{x}$  from the design origin. This implies that  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  is the same at any two locations that are the same distance from the origin.

In the case of a **second-order model**, the moments that affect rotatability (or any variance property) are moments through order four. Necessary and sufficient conditions for rotatability are as follows:

1. All odd moments through order four are zero.
  2. The ratio of moments  $[iiii]/[iijj] = 3$  ( $i \neq j$ ).
- (8.14)

The conditions are not only simple, but relatively easy to achieve, particularly with a CCD.

**8.4.3 Rotatability and the CCD**

The conditions given in Equation 8.14 are achieved, at least approximately, by several classes of designs. In the case of the CCD, rotatability is achieved by making a proper choice of  $\alpha$ , the axial distance. Condition 1 above will hold as long as the factorial portion is a full  $2^k$  or a fraction with resolution V or higher. The balance between +1 and -1 in the factorial columns and the orthogonality among certain columns in the matrix  $\mathbf{X}$  for the CCD will result in all odd moments being zero. For condition 2, one merely seeks  $\alpha$  for which

$$\frac{[iiii]}{[ijij]} = \frac{F + 2\alpha^4}{F} = 3$$

which results in

$$\alpha = \sqrt[4]{F} \tag{8.15}$$

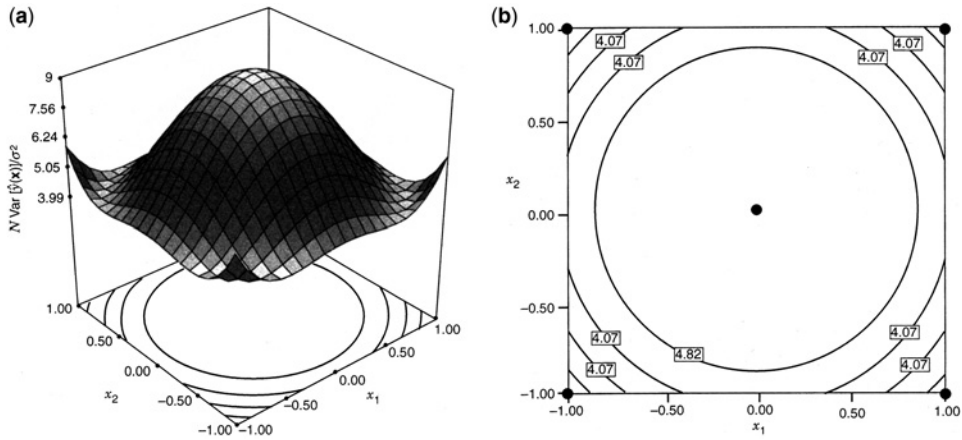
where, of course,  $F$  is the number of factorial points ( $F = 2^k$  if it is a full factorial). It is important to note that rotatability is achieved by using  $\alpha$  as in Equation 8.15 *regardless of the number of center runs*. Table 8.5 gives value of  $\alpha$  for a rotatable design for various numbers of design variables.

Note that for  $k=2$  and  $k=4$  the rotatable CCD contains 8 and 24 points (apart from center runs), respectively, that are equidistant from the design center. For  $k=3$ , the value  $\alpha = 1.682$  corresponds to the CCD used in the breadwrapper example in Section 8.4.1. For  $k=2, 3$ , and 4 the rotatable CCD is either exactly or very nearly a spherical design; that is, all points (apart from center runs) are exactly (or approximately for  $k=3$ ) a distance  $\sqrt{k}$  from the design center.

**Center Runs for the Rotatable CCD** The property of rotatability is an attempt at producing **stability**, in the sense of constant SPV( $\mathbf{x}$ ) on spheres. However, the presence of a rotatable design does not imply stability throughout the design region. In fact it turns out that a spherical design (all points on a common radius) used for fitting a second-order model has an infinite SPV( $\mathbf{x}$ ), since the design is **singular**; that is,  $\mathbf{X}'\mathbf{X}$  is a singular matrix [see Box and Hunter (1957), Box and Draper (1975), and Myers (1976)]. The use of center runs can provide reasonable *stability* of SPV( $\mathbf{x}$ ) in the design region; as a result, some center runs

**TABLE 8.5 Values of  $\alpha$  for a Rotatable Central Composite Design**

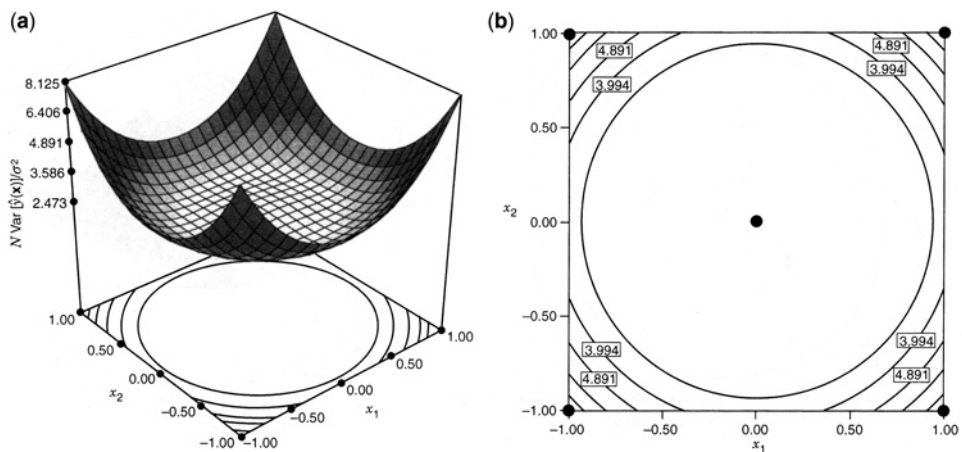
| $k$                    | $F$ | $N$         | $\alpha$ |
|------------------------|-----|-------------|----------|
| 2                      | 4   | $8 + n_c$   | 1.414    |
| 3                      | 8   | $14 + n_c$  | 1.682    |
| 4                      | 16  | $24 + n_c$  | 2.000    |
| 5                      | 32  | $42 + n_c$  | 2.378    |
| 5 ( $\frac{1}{2}$ rep) | 16  | $26 + n_c$  | 2.000    |
| 6                      | 64  | $76 + n_c$  | 2.828    |
| 6 ( $\frac{1}{2}$ rep) | 32  | $44 + n_c$  | 2.378    |
| 7                      | 128 | $142 + n_c$ | 3.364    |
| 7 ( $\frac{1}{2}$ rep) | 64  | $78 + n_c$  | 2.828    |



**Figure 8.8** Scaled prediction variance  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  for  $k=2$  CCD,  $\alpha = \sqrt{2}$ ,  $n_c = 1$ . (a) Response surface. (b) Contour plot.

for a rotatable CCD are very beneficial. The use of a rotatable or near-rotatable CCD with only a small number of center runs is not a good practice. An illustration of this point is given through Figs. 8.8 and 8.9. Figure 8.8 shows contours of SPV( $\mathbf{x}$ ) for a  $k=2$  CCD ( $\alpha = \sqrt{2}$ ) and one center run. (For zero center runs the design is singular.) Figure 8.9 gives the contours with  $\alpha = \sqrt{2}$  and five center runs. Note that the design in Fig. 8.9 is preferable. Also note that the criterion involves weighting by  $N$ , which means that the design in Fig. 8.9 has a larger weight. In spite of this, the  $n_c = 1$  design has a scaled prediction variance at the design center that is 3.5 times as large as that of the  $n_c = 5$  design.

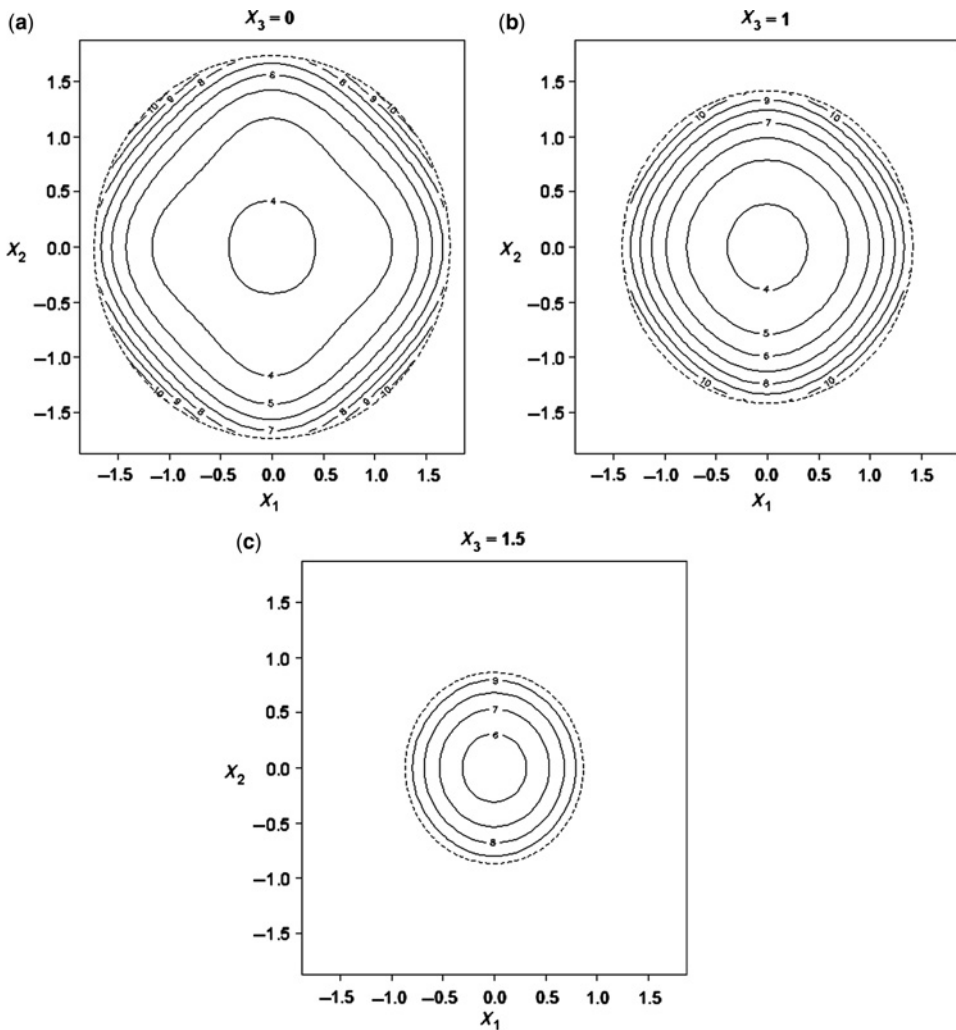
Guidelines regarding the use of center runs with the CCD will be given in the next section; suffice it to say at this point that spherical or nearly spherical designs require three to five center runs in order to avoid a severe imbalance in SPV( $\mathbf{x}$ ) through the design



**Figure 8.9** Scaled prediction variance  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  for  $k=2$  CCD,  $\alpha = \sqrt{2}$ ,  $n_c = 5$ . (a) Response surface. (b) Contour plot.

region. The message communicated by Figs. 8.8 and 8.9 extends to larger values of  $k$ . Draper (1982), Giovannitti-Jensen and Myers (1989), and Myers et al. (1992b) supply information regarding center runs in the use of the CCD.

Achieving near rotatability for a design can have the desired effect of producing good stability of prediction variance throughout the design region, even if the conditions for rotatability were not met exactly. Consider the experiment for finding the optimal conditions for storing bovine semen given in Example 6.2 with data shown in Table 6.4. In this example, three factors were considered, and the design selected was a CCD with four center runs and  $\alpha = 2$ . Figure 8.10 shows three slices of the spherical design region, based on fixed values of factor  $X_3$ . The dashed line on the plot shows the region of the design space that lies within the sphere of radius  $\sqrt{3}$  from the center. Note that this changes depending on



**Figure 8.10** Contour plot of scaled prediction variance for  $k = 3$  CCD,  $\alpha = 2$ ,  $n_c = 4$  for Example 6.2. (a)  $X_3 = 0$ , (b)  $X_3 = 1$ , (c)  $X_3 = 1.5$ .

the value of  $X_3$  considered. In this case the axial design points lie outside of the spherical region, still within the region of operability, but outside the region of interest.

Although the axial distance was not 1.682 as recommended for a rotatable design, we can see from Fig. 8.10 that the contours for this design are nearly circular, and hence the design is very close to rotatable. This illustrates that the changes in the axial distance affect rotatability relatively slowly.

**How Important is Rotatability?** The rotatable CCD plays an important role in RSM, from both a historical and an operational point of view. However, it is important for the analyst to understand that it is not necessary to have exact rotatability in a second-order design. In fact:

If the desired region of the design is spherical, the CCD that is most effective from a variance point of view is to use  $\alpha = \sqrt{k}$  and three to five center runs.

This design is not necessarily rotatable, but is near-rotatable. The recommendation is based on the stability and size of  $SPV(\mathbf{x})$  in the spherical design region. For example, for  $k = 3, \alpha = \sqrt{3}$  does not produce a rotatable design. However, the loss in rotatability is actually trivial, and the larger value of  $\alpha$  (then the rotatable value  $\alpha = 1.682$ ), results in a design that is slightly preferable based on improved stability and minimizing the average and worst-case prediction variance. Numerical evidence regarding this recommendation will be given later in this chapter and in Chapter 9. The reader is also referred to Box and Draper (1987), Khuri and Cornell (1996), Lucas (1976), Giovannitti-Jensen and Myers (1989), and Myers et al. (1992b) for information regarding the practical use of the CCD.

#### 8.4.4 More on Prediction Variance—Scaled, Unscaled, and Estimated

The importance of the **prediction variance** or the **variance of a predicted value**, as defined by

$$PV(\mathbf{x}) = \text{Var}[\hat{y}(\mathbf{x})] = \sigma^2 \mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)}$$

has importance in several phases of the design and analysis of an experiment. The presence of the unknown quantity,  $\sigma^2$ , needs special consideration and different treatment depending on whether data have been collected.

During the design selection stage, no value of  $\sigma^2$  is available. In addition, its value is not informative to the quality of the designs being compared or assessed. Hence, the scaled prediction variance,  $SPV(\mathbf{x}) = N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ , as defined in Equation 8.10 is a common choice in this stage. In their epic paper in 1957, Box and Hunter used it in graphics to compare competing designs with considerable work attached to the central composite design and other spherical designs. In fact the term “scaled” prediction was not used in their 1957 paper. Rather they used the term “information” to describe the concept. We now describe the origin and reasonableness of this term.

Cost and general design efficiency are important, and it is often helpful to weigh these relative to each other. As we discussed in Section 8.3.4, the SPV works through the use of the *information matrix*,  $\mathbf{X}'\mathbf{X}/N$ , also called the *moment matrix*. This quantity considers quality on a per *observation basis*. Thus many reasonable comparisons between designs



take cost (choice of  $N$ ) into account, and balance cost with efficiency. For the SPV can be written

$$\begin{aligned}\text{SPV} &= \mathbf{x}^{(m)'} \left( \frac{\mathbf{X}'\mathbf{X}}{N} \right)^{-1} \mathbf{x}^{(m)} \\ &= N\mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)}\end{aligned}$$

In fact, as will be demonstrated in Chapter 9, an entire spectrum of optimality criteria, such as  $D$ -optimality,  $I$ -optimality,  $G$ -optimality, and  $A$ -optimality, involve criteria that are norms on  $\mathbf{X}'\mathbf{X}/N$ , and hence use scaling as in the case of SPV.

However, the assumption that the cost of the design is proportional to its size,  $N$ , may not be appropriate for all situations. For example, consider the situation where setting up the general conditions for the experiment are more expensive than the incremental cost of performing runs of the experiment. In this case, we are more likely to want to run an experiment as large as possible given practical constraints, to maximize the benefit of our set-up. Here we might choose to compare designs of different sizes without the penalty term of  $N$  in SPV. Here the quantity, called **unscaled prediction variance**

$$\text{UPV} = \text{Var}[\hat{y}(\mathbf{x})] / \sigma^2 = \mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)}$$

may be more appropriate. There is considerable discussion among statisticians about the relative merits of the quantities, SPV and UPV. On the theoretical grounds outlined above, SPV is a useful metric for assessing the quality of prediction in the design region taking account of the size of the design. If cost as quantified by the size of the design is not a consideration for selecting between designs, or if the incremental cost of increasing the design size is not appropriately approximated by the scaling  $N$ , then UPV is a good alternative. It gives an absolute measure of the precision of a design. From a practical perspective, the SPV approach is similar to the Derringer-Suich (1980) desirability function approach described in Chapter 7, which seeks to balance good prediction quality with cost considerations. If it is more beneficial to consider these aspects of a design separately, then the UPV may be preferred. See Anderson-Cook et al. (2009a,b) with discussion for different perspectives on the merits of scaling or not.

To summarize, we can use SPV or UPV during the design selection phase. The division by  $\sigma^2$  means that these quantities are a function of only the design matrix and do not require data to have been collected. However, there are certainly many applications of the use of prediction variance in which the scaling of  $N/\sigma^2$  for the SPV or  $1/\sigma^2$  for the UPV should not be used. After the experiment has been run, and data have been collected, the estimate MSE for  $\sigma^2$  should be multiplied by the unscaled quantity  $\mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)}$  with the result being the **estimated prediction variance**,

$$\text{EPV} = \text{MSE } \mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)}$$

Here the square root of this quantity is the familiar standard error of the estimated mean,  $\hat{y}(\mathbf{x})$ , at a given location in the design space. Students are taught that at a location,  $\mathbf{x}_0$  (in the model space), the  $100(1 - \alpha)\%$  confidence interval on a mean response is given by the familiar expression

$$\hat{y}(\mathbf{x}_0) \pm t_{\alpha/2, \text{df}(\text{error})} \sqrt{\text{MSE } \mathbf{x}_0^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0^{(m)}}$$

Of course it would make no sense to scale the quantity inside the square root by a factor of  $N$ . In fact, the confidence region would not be correct with such scaling. Plots of this quantity for the variance at  $\mathbf{x}_0$  in the design region are available in several commercial software packages including Design-Expert. For example, plots in Figs. 2.13a and b for the data of Table 2.8 show illustrations of these plots.

From the foregoing, it should be surmised that if an experimental design has been implemented with the data collected, the EPV should be applied for inferences to be drawn concerning the use of the fitted model. In fact, under these conditions, the cost involved and the efficiency (compared to other designs) are no longer issues. The questions that are now being answered center around “How good is the fitted model that was generated from this design?” Thus there is no need for scaling, and using the estimated value of  $\sigma^2$  provides information about prediction that is specific to the experiment.

### 8.4.5 The Face-Centered Cube in Cuboidal Regions

In Section 8.4.2, recommendations were given for the choice of  $\alpha$  and  $n_c$  for a central composite design in a spherical region. However, there are many practical situations in which the scientist or engineer specifies ranges on the design variables, and these ranges are strict. That is, the region of interest and the region of operability are the same, and the obvious region for the design is a square ( $k = 2$ ), cube ( $k = 3$ ), or hypercube ( $k \geq 4$ ).

For example, in an experimental study designed to study organism growth, the design variables and their ranges are percent glucose [2%, 4%], percent yeast [0.4%, 0.6%], and time [30 hr, 60 hr]. Suppose it is of interest to build a second-order response surface model and the biologist is interested in predicting growth of the organism inside and on the perimeter of the cuboidal region produced by the cube. In addition, for biological reasons, one cannot experiment outside the cube, though experimentation at the extremes in the region is permissible and, in fact, desirable. This scenario, which occurs frequently in many scientific areas, suggests a central composite design in which the eight corners of the cube are centered and scaled to  $(+1, +1, +1)$  and  $\alpha = 1$ . The final design is given (in coded form) by

$$\mathbf{D} = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \left[ \begin{array}{ccc} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array} \right] \end{array}$$

where the  $(\mathbf{0}, \mathbf{0}, \mathbf{0})$  at the design center indicates a vector of  $n_c$  center runs. Figure 8.11 shows the design, often called the **face-centered cube** (or FCD) because the axial points

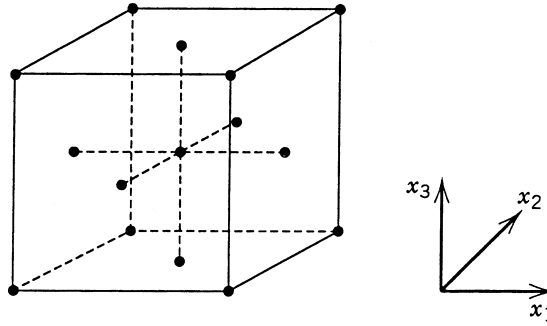


Figure 8.11 Face-centered cube (FCD with  $\alpha = 1.0$ ) for  $k = 3$ .

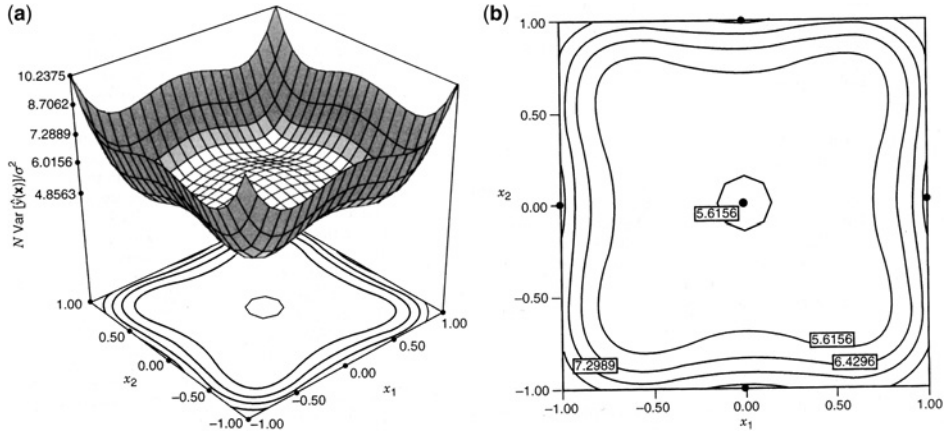
occur at the centers of the faces, rather than outside the faces as in the case of a spherical region.

In the case of a **cuboidal design region**, the face-centered cube is an effective second-order design. When one encounters what is a natural cuboidal region, it is important that the points be pushed to the extreme of the experimental region. This results in the most attractive distribution of  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ . It is important for the region to be covered in a symmetric fashion. The face-centered cube accomplishes this. Of course, the design is not rotatable. However, rotatability or near-rotatability is not an important priority when the region of interest is clearly cuboidal. It is a useful option that comes from spherical or near-spherical designs; these designs are certainly appropriate for spherical regions of interest or regions of operability, but are less appropriate with cuboidal regions.

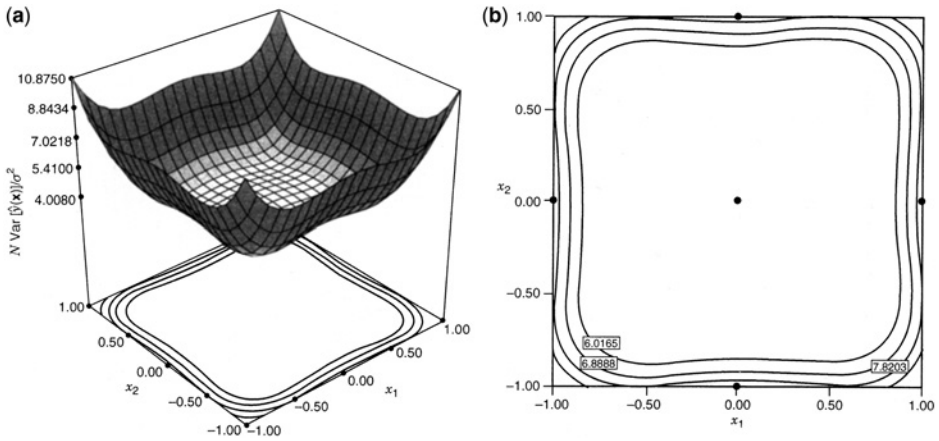
The face-centered cube is a useful design for any number of design variables. Again, a resolution V fraction is used for the factorial portion. The recommendation for center runs is quite different from that of the spherical designs. In the case of spherical designs, center runs are a necessity in order to achieve a reasonable distribution of  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ , with  $n_c = 3-5$  giving good results. In the *cuboidal* case (i.e., with  $\alpha = 1.0$ ), one or two center runs are sufficient to produce reasonable stability of  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ . The sensitivity of  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  to the number of center runs for the spherical design is seen in Figs. 8.8 and 8.9. The insensitivity to center runs for the face-centered cube is illustrated in Figs. 8.12, 8.13, and 8.14. Contours of values of  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  are given for  $n_c = 0$ ,  $n_c = 1$ , and  $n_c = 2$  for the  $k = 3$  face-centered cube when  $x_3$  is fixed at 0. It is clear that many center runs are not needed to stabilize the prediction variance. In fact, one center run is quite sufficient for stability, though  $n_c = 2$  is slightly preferable. No further improvement is achieved beyond  $n_c = 2$ .

Though much has been said here about the effect of center runs on the scaled prediction variance  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  for both spherical and cuboidal designs, it should be noted that multiple center runs or replication of exterior points may, in many cases, be desirable in order to have a sufficient number of degrees of freedom for pure error.

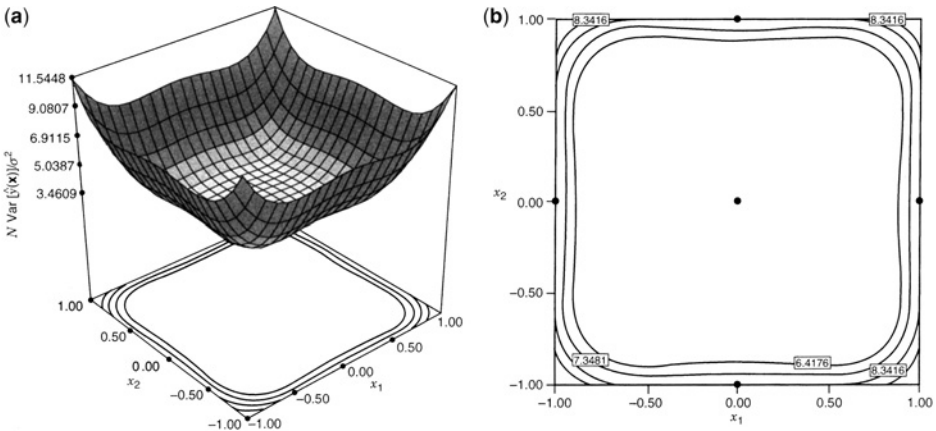
**Other Three-Level Designs on Cubes** There are other approaches to obtaining three-level designs on cubes for fitting the second-order model. While the complete  $3^k$  requires too many runs, an experimenter can either use a suitable fraction of the  $3^k$  or construct a computer-generated design based on some **alphabetic optimality criterion** (see Chapter 9). See Hoke (1974) and Mitchell and Bayne (1978) for examples of these approaches.



**Figure 8.12** Scaled prediction variance  $N \text{Var}[\hat{y}(x)]/\sigma^2$ :  $\alpha = 1.0$ ,  $k = 3$ ,  $n_c = 0$  with  $x_2 = 0$ . (a) Response surface. (b) Contour plot.



**Figure 8.13** Scaled prediction variance  $N \text{Var}[\hat{y}(x)]/\sigma^2$ :  $\alpha = 1.0$ ,  $k = 3$ ,  $n_c = 1$  with  $x_3 = 0$ . (a) Response surface. (b) Contour plot.



**Figure 8.14** Scaled prediction variance  $N \text{Var}[\hat{y}(x)]/\sigma^2$ :  $\alpha = 1.0$ ,  $k = 3$ ,  $n_c = 2$  with  $x_3 = 0$ . (a) Response surface. (b) Contour plot.

Recently, Morris (2000) has proposed a new class of second-order three level designs called augmented pairs designs, which are based on minimax and maximin distance criteria. Li et al. (2009) and Park et al. (2005) consider various designs on the cuboidal region for different numbers of factors, with graphical summaries to compare their prediction quality.

#### 8.4.6 Choosing between Spherical and Cuboidal Regions

As we have demonstrated in the previous sections, the central composite design is well suited for use in either a spherical or cuboidal design region, when the experimenter has made suitable choices about values of  $\alpha$  and  $n_c$  to ensure good design performance.

The face center or cuboidal CCD is appropriate when a cuboidal region of interest and region on operability are suggested for the experiment. In these cases the application should suggest ranges of the factors where the **corners** of the region are anticipated to produce meaningful responses and where it is natural to place factorial points. Then the question regarding the design region should hinge on whether axial points outside the ranges are scientifically permissible and should be included in the region of interest. For example, consider Fig. 8.15. The shaded area forms the cube, but after some deliberation it is determined not only that the nonshaded area is within the region of operability but also that the researcher is interested in predicting response in the unshaded area as well as the shaded area. The region of interest is a sphere circumscribed around a cube. A spherical design (CCD with  $\alpha = \sqrt{k}$ ) is certainly appropriate.

A second situation can occur, similar to that shown in Fig. 8.16. Ranges are chosen on design variable, but as the planning of the experiment evolves, it is determined that several (and perhaps all) of the vertices on the cube defined by the ranges may not be scientifically tractible; that is, they are outside both regions of operability, and the region of interest (e.g., in the case of a food product, high levels of flour, shortening, and baking time) is known to produce an unacceptable product. As a result, the corners, shaded in Fig. 8.16, are shaved off, and the design region is formed from the unshaded region. Here, the sphere is inscribed inside the initial cuboidal region formed from the selection of ranges, and the factorial point locations are moved towards the center of the design space to select combinations of factors to yield permissible results. Once this reconfiguring has been done, then selecting a scaling of the factors to more standard scales should be performed.

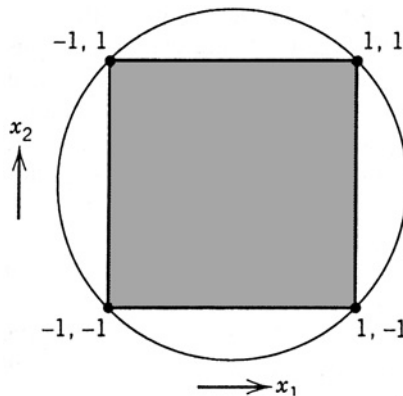


Figure 8.15 Spherical region (sphere circumscribed).

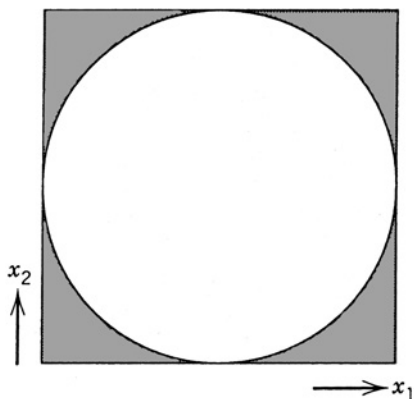


Figure 8.16 Spherical region (sphere inscribed).

Sometimes the region of operability and the region of interest are not identical. In the case where it is possible to operate slightly outside of a cuboidal region, but primary interest lies in the cube, it is possible to adjust the axial distance of the CCD to take advantage of this. Li et al. (2009) show how “*practical  $\alpha$  values*” [first suggested by Oehlert and Whitcomb (2002)] can be selected to substantially improve prediction performance in the cuboidal region. These practical  $\alpha$  values are larger than 1, and hence fall outside of the region of interest.

For larger numbers of factors, another class of potential designs to consider are called **minimum-run resolution (MinRes) V designs**. They are equireplicated two-level irregular fractions of resolution V. An equireplicated design is one in which each factor has an equal number of high and low levels. For a complete description of the construction of these designs, see Oehlert and Whitcomb (2002).

It is important to understand that in many situations the region of interest, design region, and perhaps even the region of operability are not clear cut. It is often difficult enough to get a commitment from a scientist or engineer on what are the interesting and permissible ranges of the factors. Frequently there is just insufficient knowledge a priori to know the limits of where meaningful responses can be obtained. Mistakes are often made, and adjustments using this updated knowledge are adopted into future experiments. Confusion regarding type of design or the precise region of operability should never be an excuse for not using designed experiments. Using a spherical region when it is more naturally cuboidal, for example, will still provide important information that will, among other things, lead to more educated selection of regions for future experiments.

Whether the CCD is used in a spherical or cuboidal region, it is an efficient design that is ideal for sequential experimentation and allows a reasonable amount of information for testing lack of fit while not involving an unusually large number of design points. The design accommodates a spherical region with five levels of each factor and a choice of  $\alpha = \sqrt{k}$ . The design can be a three-level design to accommodate a cuboidal region with the choice of  $\alpha = 1$ . In the spherical case, the design is either rotatable or very near rotatable, and three to five center runs should be used. In the cuboidal case, one or two center runs will suffice. Though the cuboidal and spherical designs are natural choices depending on the region of interest, the reader should not get the impression that the use of the CCD

must be confined to those choices of  $\alpha$ . There will be practical situations for which  $\alpha = 1$  or  $\alpha = \sqrt{k}$  cannot be used. The CCD should not be ruled out in these situations.

The natural competitor for the CCD is, of course, the three-level factorial, that is, the  $3^k$  factorial. The  $3^2$  design is, in fact, the face-centered cube and is thus a CCD. But when  $k$  becomes moderately large, the  $3^k$  factorial design involves an excessive number of design points. In the case of three factors and a natural cuboidal region, the  $3^3$  factorial is an efficient design if the researcher can afford to use the required 27 design points. However, for  $k > 3$  the number of design points for the  $3^k$  is usually considered impractical for most applications. A comparison of the  $3^3$  with the face-centered cube and other designs with consideration of prediction variance is given in Chapter 9.

### 8.4.7 The Box–Behnken Design

Box and Behnken (1960) developed a family of efficient **three-level designs** for fitting second-order response surfaces. The methodology for design construction is interesting and quite creative. The class of designs is based on the construction of **balanced incomplete block designs**. For example, a balanced incomplete block design with three treatments and three blocks is given by

|         | Treatment |   |   |
|---------|-----------|---|---|
|         | 1         | 2 | 3 |
| Block 1 | X         | X |   |
| Block 2 | X         |   | X |
| Block 3 |           | X | X |

The pairing together of treatments 1 and 2 symbolically implies, in the response surface setting, that design variables  $x_1$  and  $x_2$  are paired together in a  $2^2$  factorial (scaling  $\pm 1$ ) while  $x_3$  remains fixed at the center ( $x_3 = 0$ ). The same applies for blocks 2 and 3, with a  $2^2$  factorial being represented by each pair of treatments while the third factor remains fixed at 0. As a result, the  $k = 3$  Box–Behnken design is given by

$$\mathbf{D} = \begin{matrix} & x_1 & x_2 & x_3 \\ \begin{bmatrix} -1 & -1 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & -1 \\ -1 & 0 & 1 \\ 1 & 0 & -1 \\ 1 & 0 & 1 \\ 0 & -1 & -1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \end{matrix}$$

The last row in the design matrix implies a *vector* of center runs. In the case of  $k=4$  the same methodology applies. Each pair of factors are linked in a  $2^2$  factorial. The design matrix is given by

$$\mathbf{D} = \begin{matrix} & x_1 & x_2 & x_3 & x_4 \\ \begin{bmatrix} -1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} & & & & \end{matrix} \tag{8.16}$$

Note that the Box–Behnken design (BBD) is quite comparable in number of design points to the CCD for  $k=3$  and  $k=4$ . (There is no BBD for  $k=2$ .) For  $k=3$ , the CCD contains  $14 + n_c$  runs while the BBD contains  $12 + n_c$  runs. For  $k=4$  the CCD and BBD both contain  $24 + n_c$  design points. For  $k=5$  we have the following  $40 + n_c$  experimental runs:

$$\mathbf{D} = \begin{matrix} & x_1 & x_2 & x_3 & x_4 & x_5 \\ \begin{bmatrix} \pm 1 & \pm 1 & 0 & 0 & 0 \\ \pm 1 & 0 & \pm 1 & 0 & 0 \\ \pm 1 & 0 & 0 & \pm 1 & 0 \\ \pm 1 & 0 & 0 & 0 & \pm 1 \\ 0 & \pm 1 & \pm 1 & 0 & 0 \\ 0 & \pm 1 & 0 & \pm 1 & 0 \\ 0 & \pm 1 & 0 & 0 & \pm 1 \\ 0 & 0 & \pm 1 & \pm 1 & 0 \\ 0 & 0 & \pm 1 & 0 & \pm 1 \\ 0 & 0 & 0 & \pm 1 & \pm 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} & & & & & \end{matrix} \tag{8.17}$$



The CCD for  $k = 5$  involves  $26 + n_c$  runs when the  $\frac{1}{2}$  fraction is used in the factorial portion. When the full factorial is used, the CCD makes use of  $42 + n_c$  runs. Each row in the BBD symbolizes four design points with  $(\pm 1, \pm 1)$  representing a  $2^2$  factorial.

For  $k = 6$  the construction of the design is based on partially balanced incomplete block designs. Thus, each treatment does not occur with every other treatment the same number of times. In the Box–Behnken construction, this means that each factor does not occur in a two-level factorial structure the same number of times with every factor. The design matrix, involving  $N = 48 + n_c$  runs, is given by

$$\mathbf{D} = \begin{matrix} & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ \begin{bmatrix} \pm 1 & \pm 1 & 0 & \pm 1 & 0 & 0 \\ 0 & \pm 1 & \pm 1 & 0 & \pm 1 & 0 \\ 0 & 0 & \pm 1 & \pm 1 & 0 & \pm 1 \\ \pm 1 & 0 & 0 & \pm 1 & \pm 1 & 0 \\ 0 & \pm 1 & 0 & 0 & \pm 1 & \pm 1 \\ \pm 1 & 0 & \pm 1 & 0 & 0 & \pm 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} & & & & & & \end{matrix} \quad (8.18)$$

The CCD requires  $44 + n_c$  runs when the one-half fraction is used in the factorial portion.

Unlike the  $k = 3, 4,$  and  $5$  cases, the factorial structures in the BBD for  $k = 6$  are  $2^3$  factorials involving three factors. Thus each row of the above matrix involves eight design points.

For  $k = 7$ , the factorial structures again involve combinations of three design factors. The  $N = 56 + n_c$  design runs are given by the following design matrix:

$$\mathbf{D} = \begin{matrix} & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 \\ \begin{bmatrix} 0 & 0 & 0 & \pm 1 & \pm 1 & \pm 1 & 0 \\ \pm 1 & 0 & 0 & 0 & 0 & \pm 1 & \pm 1 \\ 0 & \pm 1 & 0 & 0 & \pm 1 & 0 & \pm 1 \\ \pm 1 & \pm 1 & 0 & \pm 1 & 0 & 0 & 0 \\ 0 & 0 & \pm 1 & \pm 1 & 0 & 0 & \pm 1 \\ \pm 1 & 0 & \pm 1 & 0 & \pm 1 & 0 & 0 \\ 0 & \pm 1 & \pm 1 & 0 & 0 & \pm 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} & & & & & & \end{matrix} \quad (8.19)$$

**Characteristics of the Box–Behnken Design** In many scientific studies that require RSM, researchers are inclined to require three evenly spaced levels. Thus, the Box–Behnken design is an efficient option and indeed an important alternative to the central composite design. As we can observe from the sample sizes, there is sufficient information available for testing lack of fit. For example, for  $k = 6$  the use of, say,  $n_c = 5$  would allow four degrees of freedom for pure error and 21 degrees of freedom for lack of fit. The Box–Behnken design does not substantially deviate from rotatability, and, in fact, for  $k = 4$  and  $k = 7$  the design is exactly rotatable. Verification of rotatability in both cases should be quite simple for the reader. From the design matrix in Equation 8.16, it is easy to see that all odd moments are zero. This is a result of the prominence of the sets of  $2^2$  factorial arrays. For the second condition required for rotatability, note from Equation 8.16 that  $(iii) = \frac{12}{24}$  and

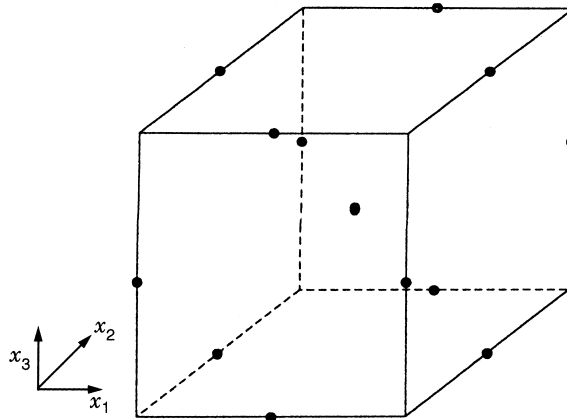


Figure 8.17 The  $k = 3$  BBD with a center point.

$(iijj) = \frac{4}{24}$ , for  $i \neq j$ . As a result,  $(iiii)/(iijj) = 3$ . An investigation of Equation 8.19 suggests that the rotatability conditions hold for the  $k = 7$  BBD as well.

Another important characteristic of the BBD is that it is a **spherical design**. Note, for example, in the  $k = 3$  case that all of the points are so-called “edge points” (i.e., points that are on the edges of the cube); in this case, all edge points are a distance  $\sqrt{2}$  from the design center. There are no factorial points or face points. Figure 8.17 displays the BBD for  $k = 3$ . Contrast this design to that depicted in Fig. 8.11 with the face-centered cube. The latter displays face points and factorial points that are not the same distance from the design center, though they all reside on the cube and provide good coverage of the cube. The BBD involves all edge points, but the entire cube is not covered. In fact, there are no points on the corner of the cube or even a distance  $\sqrt{3}$  from the design center. The analyst should not view the lack of coverage of the cube as a reason not to use the BBD. It is not meant to be a cuboidal design. However, the use of the BBD should be confined to situations in which one is *not interested in predicting response at the extremes*; that is, at the corners of the cube. A quick inspection of the design matrices in Equations 8.16–8.19 further underscores the fact that the BBD is not a cuboidal design. For example, for  $k = 7$ , the design points are at a radius of  $\sqrt{3}$ , which is considerably smaller than the radius  $\sqrt{7}$  of the corner of the cube. If three levels are required and coverage of the cube is necessary, one should use a face-centered cube rather than the BBD.

The spherical nature of the BBD, combined with the fact that the designs are rotatable or near-rotatable, suggests that ample center runs should be used. In fact, for  $k = 4$  and 7, center runs are necessary to avoid singularity. The use of three to five center runs is recommended for the BBD.

**Example 8.5 A BBD for Optimizing Viscosity** One step in the production of a particular polyamide resin is the addition of amines. It was felt that the manner of addition has a profound effect on the molecular weight distribution of the resin. Three variables are thought to play a major role: temperature at the time of addition ( $x_1$ , °C), agitation ( $x_2$ , rpm), and rate of addition ( $x_3$ ,  $\text{min}^{-1}$ ). Because it was difficult to physically set the

**TABLE 8.6 Data for Resin Viscosity in Example 8.5 Using Box–Behnken Design**

| Level          | Temperature | Agitation | Rate  | $x_1$ | $x_2$ | $x_3$ |
|----------------|-------------|-----------|-------|-------|-------|-------|
| High           | 200         | 10.0      | 25    | +1    | +1    | +1    |
| Center         | 175         | 7.5       | 20    | 0     | 0     | 0     |
| Low            | 150         | 5.0       | 15    | -1    | -1    | -1    |
| Standard Order |             | $x_1$     | $x_2$ | $x_3$ | $y$   |       |
| 1              |             | -1        | -1    | 0     | 53    |       |
| 2              |             | +1        | -1    | 0     | 58    |       |
| 3              |             | -1        | +1    | 0     | 59    |       |
| 4              |             | +1        | +1    | 0     | 56    |       |
| 5              |             | -1        | 0     | -1    | 64    |       |
| 6              |             | +1        | 0     | -1    | 45    |       |
| 7              |             | -1        | 0     | +1    | 35    |       |
| 8              |             | +1        | 0     | +1    | 60    |       |
| 9              |             | 0         | -1    | -1    | 59    |       |
| 10             |             | 0         | +1    | -1    | 64    |       |
| 11             |             | 0         | -1    | +1    | 53    |       |
| 12             |             | 0         | +1    | +1    | 65    |       |
| 13             |             | 0         | 0     | 0     | 65    |       |
| 14             |             | 0         | 0     | 0     | 59    |       |
| 15             |             | 0         | 0     | 0     | 62    |       |

levels of addition and agitation, three levels were chosen and a BBD was used. The viscosity of the resin was recorded as an indirect measure of molecular weight. The data, including natural levels, design, and response values, are shown in Table 8.6. Figure 8.18 shows an analysis using the package JMP. Three contour plots of constant viscosity (shown in Fig. 8.19) are given in order to illustrate what combination of the factors produce high- and low-molecular-weight resins. Agitation was fixed at the low, medium, and high levels. It is clear that high-molecular-weight resins are produced with low values of agitation, whereas low-molecular-weight resins are potentially available when one uses high values of agitation. In fact, the extremes of low rates, low temperature, and high agitation produces low-molecular-weight resins, whereas low agitation, low temperature, and high rate results in high-molecular-weight resins.

#### 8.4.8 Definitive Screening Designs for Fitting Second-Order Models

In Section 8.3.4 we introduced the class of Definitive Screening Designs (DSD) as models suitable for fitting first-order models with interactions. There, we noted that the use of three factor levels allowed for the separate estimation of all of the quadratic terms, even though there might not be sufficient degrees of freedom to fit the full second-order model. We now consider these designs for fitting second-order models, with a strong cautionary note that the experimenter is relying heavily on effect sparsity in order for this to be a successful strategy. If too many of the terms in the second-order model are active, then it may be difficult to identify which of the terms should be included in the final model, using the model selection strategies such as forward stepwise regression outlined in Section 4.9.

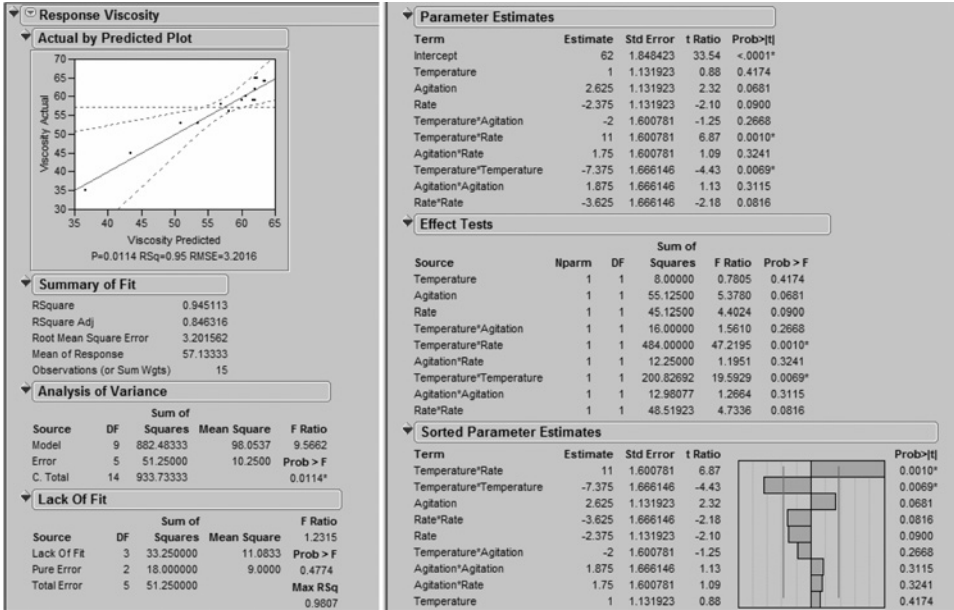


Figure 8.18 JMP analysis of resin viscosity data of Example 8.5.

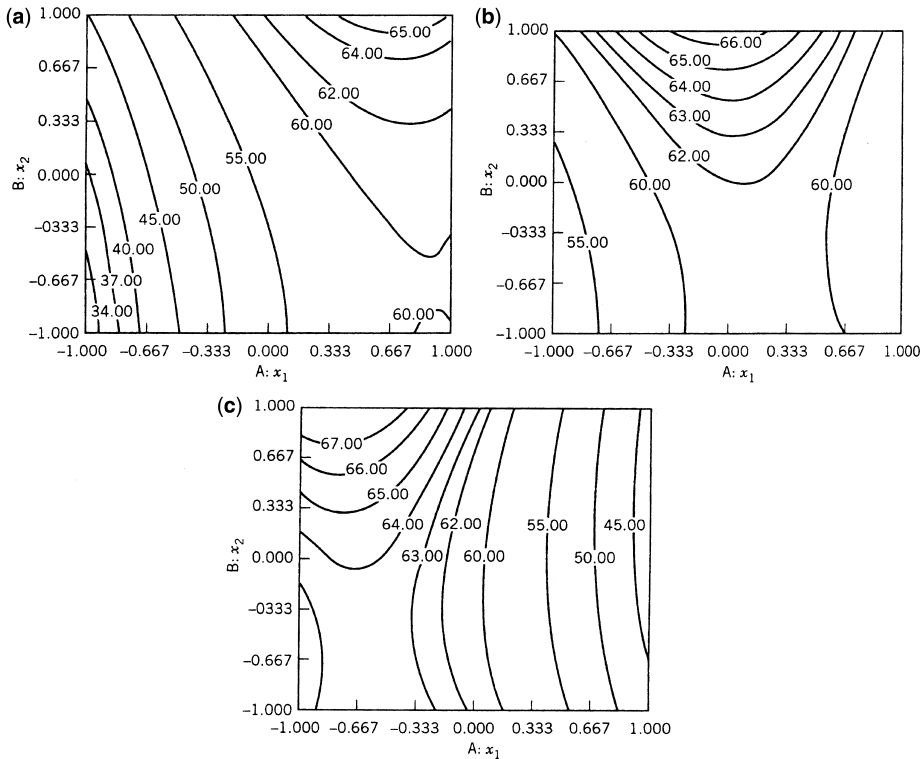


Figure 8.19 Contour plots of viscosity from Example 8.5. (a) Addition rate  $x_3 = 1$ . (b) Addition rate  $x_3 = 0$ . (c) Addition rate  $x_3 = -1$ .

An additional consideration when contemplating the use of a DSD for a combined screening experiment with the potential to fit a second-order model is whether the experimenter is confident about the design region being well-located in the region of operability. If the experimenter thinks that there is a good chance that the optimal location for the process can be included in the current design region, then the DSD becomes even more attractive. However, if the user thinks that the path of steepest ascent might be need to traverse the region of operability to get closer to the preferred location, then a simpler design, such as a factorial or fractional factorial design, may be better suited to the goals of the experiment.

In the example that follows, we illustrate how a DSD can be used to explore the impact of multiple factors on the response while still allowing a second-order model to be fit.

**Example 8.6 A Definitive Screening Design for Screening and a Potential Second-Order Model** An experimenter wishes to run a designed experiment to understand a process. Current understanding of the process suggests that three factors (A, B, and C) have an impact on the response, Y. However, several of the engineers involved with the process are concerned that there is some chance that one of several other factors (D, E, F, and G) might contribute to changes in the response. Selecting a DSD for 7 factors with 17 runs ( $2 \times 7 + 3$  center runs) allows flexibility for several potential outcomes. Table 8.7 shows the DSD for seven factors from Section 8.3.4.

If any of factors D–G are active, then there is the opportunity to detect them, while if the most-likely-to-be-active factors, A–C, dominate as expected, then we can achieve good estimation and prediction with the DSD. If the experimenter ran the design in Table 8.7 and found that a second-order model involving factors A–C was appropriate, then by examining  $(\mathbf{X}'\mathbf{X})^{-1}$  we can see how well the different terms in the model are estimated by looking at a multiple of their variance.

**TABLE 8.7 Definitive Screening Design for Seven Factors**

| Run   | A | B | C | D | E | F | G |
|-------|---|---|---|---|---|---|---|
| 1     | 0 | + | – | + | – | + | – |
| 2     | 0 | – | + | – | + | – | + |
| 3     | – | 0 | + | – | + | + | – |
| 4     | + | 0 | – | + | – | – | + |
| 5     | + | – | 0 | + | + | + | + |
| 6     | – | + | 0 | – | – | – | – |
| 7     | + | – | – | 0 | + | – | – |
| 8     | – | + | + | 0 | – | + | + |
| 9     | – | – | + | + | 0 | – | – |
| 10    | + | + | – | – | 0 | + | + |
| 11    | – | + | – | + | + | 0 | + |
| 12    | + | – | + | – | – | 0 | – |
| 13    | + | + | + | + | + | – | 0 |
| 14    | – | – | – | – | – | + | 0 |
| 15–17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

|                                  | Int    | A     | B     | C     | AB     | AC     | BC     | A <sup>2</sup> | B <sup>2</sup> | C <sup>2</sup> |
|----------------------------------|--------|-------|-------|-------|--------|--------|--------|----------------|----------------|----------------|
|                                  | 0.292  | 0     | 0     | 0     | 0.015  | 0.015  | 0.015  | -0.108         | -0.108         | -0.108         |
|                                  | 0      | 0.089 | 0.018 | 0.018 | 0      | 0      | 0      | 0              | 0              | 0              |
|                                  | 0      | 0.018 | 0.089 | 0.018 | 0      | 0      | 0      | 0              | 0              | 0              |
|                                  | 0      | 0.018 | 0.018 | 0.089 | 0      | 0      | 0      | 0              | 0              | 0              |
| $(\mathbf{X}'\mathbf{X})^{-1} =$ | 0.015  | 0     | 0     | 0     | 0.119  | -0.006 | -0.006 | 0.040          | 0.040          | -0.085         |
|                                  | 0.015  | 0     | 0     | 0     | -0.006 | 0.119  | -0.006 | 0.040          | -0.085         | 0.040          |
|                                  | 0.015  | 0     | 0     | 0     | -0.006 | -0.006 | 0.119  | -0.085         | 0.040          | 0.040          |
|                                  | -0.108 | 0     | 0     | 0     | 0.040  | 0.040  | -0.085 | 0.467          | -0.158         | -0.158         |
|                                  | -0.108 | 0     | 0     | 0     | 0.040  | -0.085 | 0.040  | -0.158         | 0.467          | -0.158         |
|                                  | -0.108 | 0     | 0     | 0     | -0.085 | 0.040  | 0.040  | -0.158         | -0.158         | 0.467          |

Note that since the design has 14 non-center runs, it is not possible for the main effects for the three factors to be orthogonal, given the structure of the DSD being constructed based on the foldover conference matrix structure. This was described in more detail in Section 8.3.4. However, the zero entries in  $(\mathbf{X}'\mathbf{X})^{-1}$  are a reminder that the main effects are orthogonal to all of the two-way interactions and quadratic terms in the model. Based on the size of the terms in the  $(\mathbf{X}'\mathbf{X})^{-1}$  matrix, it is clear that the main effects are estimated with the smallest variance, while the quadratic terms have larger associated variances. Fitting this model involves using 10 of the 17 available degrees of freedom, so there would be some opportunity to include an additional term in the model. Suppose that one unexpected main effect from factors D–G (say D) suggests that it should be included in the model, then the  $(\mathbf{X}'\mathbf{X})^{-1}$  matrix is only minimally affected.

|  | Int    | A      | B      | C     | D      | AB     | AC     | BC     | A <sup>2</sup> | B <sup>2</sup> | C <sup>2</sup> |
|--|--------|--------|--------|-------|--------|--------|--------|--------|----------------|----------------|----------------|
|  | 0.292  | 0.000  | 0.000  | 0.000 | 0.000  | 0.015  | 0.015  | 0.015  | -0.108         | -0.108         | -0.108         |
|  | 0      | 0.092  | 0.021  | 0.016 | -0.016 | 0      | 0      | 0      | 0              | 0              | 0              |
|  | 0      | 0.021  | 0.092  | 0.016 | -0.016 | 0      | 0      | 0      | 0              | 0              | 0              |
|  | 0      | 0.016  | 0.016  | 0.090 | 0.010  | 0      | 0      | 0      | 0              | 0              | 0              |
|  | 0      | -0.016 | -0.016 | 0.010 | 0.090  | 0      | 0      | 0      | 0              | 0              | 0              |
|  | 0.015  | 0      | 0      | 0     | 0      | 0.119  | -0.006 | -0.006 | 0.040          | 0.040          | -0.085         |
|  | 0.015  | 0      | 0      | 0     | 0      | -0.006 | 0.119  | -0.006 | 0.040          | -0.085         | 0.040          |
|  | 0.015  | 0      | 0      | 0     | 0      | -0.006 | -0.006 | 0.119  | -0.085         | 0.040          | 0.040          |
|  | -0.108 | 0      | 0      | 0     | 0      | 0.040  | 0.040  | -0.085 | 0.467          | -0.158         | -0.158         |
|  | -0.108 | 0      | 0      | 0     | 0      | 0.040  | -0.085 | 0.040  | -0.158         | 0.467          | -0.158         |
|  | -0.108 | 0      | 0      | 0     | 0      | -0.085 | 0.040  | 0.040  | -0.158         | -0.158         | 0.467          |

It is interesting to consider the natural design space for the DSD. When all of the seven factors are included, the points except the center runs are all located  $\sqrt{7-1}$  scaled units from the center of the design space, suggesting a spherical design region. However, when we perform model reduction and collapse the design down to just 3 or 4 factors, then some of the non-center runs do not include any zero levels, and hence a cuboidal region in the projected design space is suggested. Figure 8.20 shows the geometry of the DSD in Table 8.7 when we project the design onto the three factors, A–C.

Another possible alternative for the design, would be to select the eight-factor DSD, since with its even number of factors would allow for orthogonal estimation of all of the main effects. If the experimenter wished to stay with an overall design size of 17 runs, then it would be possible to use only a single center run. This sacrifices the ability to estimate

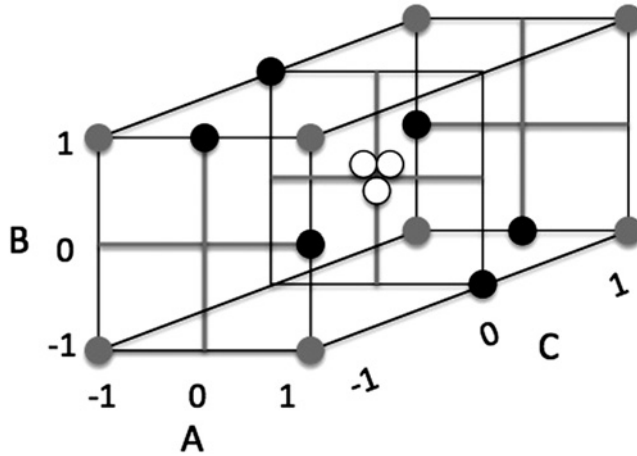


Figure 8.20 Definitive Screening Design for seven factors projected onto 3 factor space.

pure error, but the  $(\mathbf{X}'\mathbf{X})^{-1}$  matrix below when fitting the full second order model for factors A-C shows that the main effects are now all orthogonal.

| Int    | A     | B     | C     | AB     | AC     | BC     | A <sup>2</sup> | B <sup>2</sup> | C <sup>2</sup> |
|--------|-------|-------|-------|--------|--------|--------|----------------|----------------|----------------|
| 0.702  | 0     | 0     | 0     | -0.033 | -0.033 | 0.033  | -0.260         | -0.260         | -0.260         |
| 0      | 0.071 | 0     | 0     | 0      | 0      | 0      | 0              | 0              | 0              |
| 0      | 0     | 0.071 | 0     | 0      | 0      | 0      | 0              | 0              | 0              |
| 0      | 0     | 0     | 0.071 | 0      | 0      | 0      | 0              | 0              | 0              |
| -0.033 | 0     | 0     | 0     | 0.106  | -0.019 | 0.019  | -0.028         | -0.028         | 0.097          |
| -0.033 | 0     | 0     | 0     | -0.019 | 0.106  | 0.019  | -0.028         | 0.097          | -0.028         |
| 0.033  | 0     | 0     | 0     | 0.019  | 0.019  | 0.106  | -0.097         | 0.028          | 0.028          |
| -0.260 | 0     | 0     | 0     | -0.028 | -0.028 | -0.097 | 0.522          | -0.103         | -0.103         |
| -0.260 | 0     | 0     | 0     | -0.028 | 0.097  | 0.028  | -0.103         | 0.522          | -0.103         |
| -0.260 | 0     | 0     | 0     | 0.097  | -0.028 | 0.028  | -0.103         | -0.103         | 0.522          |

If we compare the performance of the DSD with a face center cube (central composite design with  $\alpha = 1$ ) for 3 factors with 3 center runs, we see that the designs both have 17 runs. The  $(\mathbf{X}'\mathbf{X})^{-1}$  matrix shows that the variance estimates of the different terms are quite comparable between the two designs.

|                                  | Int    | A     | B     | C     | AB    | AC    | BC    | A <sup>2</sup> | B <sup>2</sup> | C <sup>2</sup> |
|----------------------------------|--------|-------|-------|-------|-------|-------|-------|----------------|----------------|----------------|
| $(\mathbf{X}'\mathbf{X})^{-1} =$ | 0.183  | 0     | 0     | 0     | 0     | 0     | 0     | -0.070         | -0.070         | -0.070         |
|                                  | 0      | 0.100 | 0     | 0     | 0     | 0     | 0     | 0              | 0              | 0              |
|                                  | 0      | 0     | 0.100 | 0     | 0     | 0     | 0     | 0              | 0              | 0              |
|                                  | 0      | 0     | 0     | 0.100 | 0     | 0     | 0     | 0              | 0              | 0              |
|                                  | 0      | 0     | 0     | 0     | 0.125 | 0     | 0     | 0              | 0              | 0              |
|                                  | 0      | 0     | 0     | 0     | 0     | 0.125 | 0     | 0              | 0              | 0              |
|                                  | 0      | 0     | 0     | 0     | 0     | 0     | 0.125 | 0              | 0              | 0              |
|                                  | -0.070 | 0     | 0     | 0     | 0     | 0     | 0     | 0.373          | -0.127         | -0.127         |
|                                  | -0.070 | 0     | 0     | 0     | 0     | 0     | 0     | -0.127         | 0.373          | -0.127         |
|                                  | -0.070 | 0     | 0     | 0     | 0     | 0     | 0     | -0.127         | -0.127         | 0.373          |

the variances of the main, quadratic and two-way interaction terms are of similar magnitude. Notably, there are more zeros in the off-diagonals of the  $(\mathbf{X}'\mathbf{X})^{-1}$  matrix for the CCD, than for the seven-factor DSD. The eight-factor DSD is more comparable to the CCD, but sacrifices the ability to examine pure error. One substantial advantage of the two DSDs is their ability to evaluate potential contributions from the other factors, D–G. If the experimenter were willing to consider a design of with 19 runs, then the eight-factor DSD with three center runs might offer an ideal option, but at a slightly higher cost because of its larger size. Again, the inherent trade-offs between alternative designs suggest that the underlying assumption of the relationship between factors and response play an important role in determining which of the designs would be preferred for a particular study.

***What to Do if Too Many Terms Are Active*** Suppose that the experimenter opts to run the 17-run seven-factor DSD shown in Table 8.7. When forward selection is performed after the experimental data have been corrected, there is confusion about what terms to include in the model since a considerable number of terms appear to be of similar size, and all appear to be large relative to the estimated pure error. In this case, the experimenter may need to supplement the DSD with additional runs to clarify the results and feel more comfortable with understanding which terms are important.

A potential design to consider comes from the class of no-confounding designs described in Section 4.8.1. If the DSD in Table 8.7 is complemented with the seven-factor nonregular orthogonal design in Table 4.29, the resulting design with  $17 + 16 = 33$  runs has some appealing properties. While substantially larger than the original DSD alone, this design is still considerably smaller than having fit a seven-factor central composite design, with  $78 + n_c$  runs. In addition, the projection properties of the combined 33-run design allow the estimation of the full second-order model for all subsets of five factors chosen from the seven possible. Hence, if the engineers were right in their concerns that some of the factors D–G contribute to changes in the response, then there are good opportunities to correctly identify the appropriate model. The additional degrees of freedom available provide additional information to help disentangle which effects are active. In addition, the levels of the no-confounding design are all  $-1$  and  $+1$ , improving the quality of prediction of all of the terms in the model.

This characteristic of good performance of a two-stage design with a DSD and a follow-on no-confounding design is a good combination to consider as part of a sequential design of experiments strategy. The initial DSD may often be adequate for identifying the appropriate response surface model, but it is helpful to have a secondary plan for how to supplement this design if results are not clear.

As illustrated in this section and Section 8.3.4, the class of Definitive Screening Designs is versatile for different possible models, and it allows for efficient compact designed experiments for first-order, first-order with interaction, and second-order models. When possible, using DSDs with even numbers of factors is preferred for the additional benefit of orthogonality of the main effects with each other.

#### 8.4.9 Orthogonal Blocking in Second-Order Designs

In many RSM situations the study is too large to allow all runs to be made under homogeneous conditions. As a result, it is important and interesting to consider second-order designs that facilitate *blocking*—that is, the inclusion of block effects. It is important that the



assignment of the design points to blocks be done so as to minimize the effect on the model coefficients. The property of the experimental design that we seek is that of **orthogonal blocking**. To say that a design admits orthogonal blocking implies that the block effects in the model are orthogonal to model coefficients. A simple first-order example serves as an illustration. Consider a  $2^2$  factorial to be used for the fitting of a first-order model. Suppose we use  $I=AB$  as the defining relation to form two one-half fractions, which are placed in separate blocks. As a result, we have the model

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \delta_1 z_{i1} + \delta_2 z_{i2} + \epsilon_i \tag{8.20}$$

where  $\delta_1$  and  $\delta_2$  are **block effect** coefficients. Here  $z_{i1}$  and  $z_{i2}$  are **dummy** or indicator variables; that is,  $z_{i1} = 1$  if  $y_i$  is in block 1 and  $z_{i1} = 0$  otherwise. The variable  $z_{i2} = 1$  if  $y_i$  is in block 2 and  $z_{i2} = 0$  otherwise. The  $X$  matrix is then

$$\mathbf{X} = \begin{matrix} & x_1 & x_2 & z_1 & z_2 \\ \left. \begin{matrix} 1 & -1 & -1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & -1 & 1 & 0 & 1 \\ 1 & 1 & -1 & 0 & 1 \end{matrix} \right\} \begin{matrix} \text{Block 1} \\ \\ \\ \text{Block 2} \end{matrix} \end{matrix} \tag{8.21}$$

Note that block 1 contains the design points  $\{(1), ab\}$ , whereas block 2 contains  $\{a, b\}$ . Now, in Equation 8.21, the matrix  $\mathbf{X}$  is singular; the last two columns add to the first column. This underscores the fact that the parameters  $\delta_1$  and  $\delta_2$  are not estimable. As a result, one of the indicator variables should not be included. We eliminate the  $z_2$  column and center  $z_1$ . As a result, the model Equation 8.20 can be rewritten

$$y_i = \beta_0^* + \beta_1 x_{i1} + \beta_2 x_{i2} + \delta_1 (z_{i1} - \bar{z}_1) + \epsilon_i \tag{8.22}$$

and thus  $\mathbf{X}$  is written

$$\mathbf{X} = \begin{matrix} & \beta_1 & \beta_2 & \delta_1 \\ \left. \begin{matrix} 1 & -1 & -1 & 1/2 \\ 1 & 1 & 1 & 1/2 \\ \dots & \dots & \dots & \dots \\ 1 & -1 & 1 & -1/2 \\ 1 & 1 & -1 & -1/2 \end{matrix} \right\} \begin{matrix} \text{Block 1} \\ \\ \\ \text{Block 2} \end{matrix} \end{matrix} \tag{8.23}$$

Now, the concept of orthogonal blocking is nicely illustrated through the use of Equations 8.22 and 8.23. In this example the block effect is orthogonal to the regression coefficients if

$$\begin{aligned} \sum_{u=1}^4 (z_{u1} - \bar{z}_1) x_{1u} &= 0 \\ \sum_{u=1}^4 (z_{u1} - \bar{z}_1) x_{2u} &= 0 \end{aligned} \tag{8.24}$$

Thus, as the reader might have expected, the assignment of design points to blocks in this example provides an array in which the block factor has no effect on the regression coefficients. In fact, if one considers the least squares estimator  $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ , the regression coefficients  $\beta_0^*$ ,  $\beta_1$ , and  $\beta_2$  are estimated *as if the blocks were not in the experiment*. The variances of the estimates of  $\beta_1$  and  $\beta_2$  in this variance optimal design are not changed in the presence of blocks.

The above example serves two purposes. First, it reinforces the material in Chapter 3 regarding blocking in two-level designs. In the above illustration, the  $x_1x_2$  interaction is **confounded with blocks**. If one augments the matrix  $\mathbf{X}$  with the  $x_1x_2$  column, this confounding becomes evident. The linear effects, or linear coefficients, are independent of block effects, and this is reinforced by the fact that Equations 8.24 hold. Secondly, the example allows us a foundation for setting up conditions for orthogonal blocking for second-order designs.

Though we now move on to the second-order case, there are several exercises at the end of this chapter that deal with blocking for first-order models and first-order-plus-interaction models.

**Conditions for Orthogonal Blocking in Second-Order Designs** Consider a second-order model with  $k$  design variables and  $b$  blocks. The model, then, can be written as

$$\begin{aligned}
 y_u = & \beta_0 + \sum_{i=1}^k \beta_i x_{ui} + \sum_{i=1}^k \beta_{ii} x_{ui}^2 + \sum_{i < j=2}^k \beta_{ij} x_{ui} x_{uj} \\
 & + \sum_{m=1}^b \delta_m (z_{um} - \bar{z}_m), \quad u = 1, 2, \dots, N
 \end{aligned}
 \tag{8.25}$$

Here we have used the indicator variable again, with  $z_{mu} = 1$  if the  $u$ th observation is in the  $m$ th block. We have centered the indicator variables as in our earlier example. In addition, we have included all of the block effects in the model even though it results in a singular model. This will not affect the conditions for orthogonal blocking. From Equation 8.25, block effects are orthogonal to regression coefficients if

$$\sum_{u=1}^N x_{ui} (z_{um} - \bar{z}_m) = 0, \quad i = 1, 2, \dots, k, \quad m = 1, 2, \dots, b \tag{8.26}$$

$$\sum_{u=1}^N x_{ui} x_{uj} (z_{um} - \bar{z}_m) = 0, \quad i \neq j, \quad m = 1, 2, \dots, b \tag{8.27}$$

$$\sum_{u=1}^N x_{ui}^2 (z_{um} - \bar{z}_m) = 0, \quad i = 1, 2, \dots, k, \quad m = 1, 2, \dots, b \tag{8.28}$$

In what follows we assume that we are dealing with designs in which  $[i] = 0$  and  $[ij] = 0$  for all  $i \neq j, i, j = 1, 2, \dots, k$ . That is, the first moments and mixed second moments are all zero. Recall that these conditions hold for all central composite designs, Box–Behnken designs,

and two-level factorial designs. Now, if we consider Equations 8.26 and 8.27 for a specific value of  $m$  (i.e., for a specific block), we have

$$\sum_{\text{block } m} x_{ui} = \sum_{\text{block } m} x_{ui}x_{uj} = 0, \quad i, j = 1, 2, \dots, k, \quad i \neq j \quad (8.29)$$

But, of course, Equation 8.28 must hold for all blocks. Equation 8.29 implies, then, that in each block the first and second mixed moments are zero, implying that *each block must itself be a first-order orthogonal design* (condition 1).

Now consider Equation 8.28. For say the  $m$ th block, this equation implies that  $\sum_{u=1}^N x_{ui}^2 z_{um} = \bar{z}_m \sum_{u=1}^N x_{ui}^2$  for  $i = 1, 2, \dots, k$ . The quantity  $\bar{z}_m$  is the fraction of the total runs that are in the  $m$ th block. As a result, Equation 8.28 implies that *for each design variable, the sum of squares contribution from each block is proportional to the block size* (condition 2).

An **orthogonal block design for a second-order model** requires the following conditions:

1. Each block must consist of a first-order orthogonal design
2. For each design variable, the sum of squares contribution from each block is proportional to the size of the block.

Designs that fulfill conditions 1 and 2 above are not difficult to find. Once again, the central composite design is prominent in that respect due to flexibility involved in the choice of  $\alpha$  and  $n_c$ .

**Orthogonal Blocking in the CCD** It may be instructive to discuss an example CCD that does block orthogonally. Consider a  $k = 2$  rotatable CCD in two blocks with two center points per block. Consider the assignment of design points to blocks as follows:

| Block 1 |       | Block 2     |             |
|---------|-------|-------------|-------------|
| $x_1$   | $x_2$ | $x_1$       | $x_2$       |
| -1      | -1    | $-\sqrt{2}$ | 0           |
| -1      | 1     | $-\sqrt{2}$ | 0           |
| 1       | -1    | 0           | $-\sqrt{2}$ |
| 1       | 1     | 0           | $\sqrt{2}$  |
| 0       | 0     | 0           | 0           |
| 0       | 0     | 0           | 0           |

The above design does block orthogonally. Each block is a first-order orthogonal design; the first moments and mixed second moments are both zero for the design in each block, and thus condition 1 holds. Also, the sum of squares for each design variable is 4.0, and the block sizes are equal. Hence condition 2 regarding the proportionality relationship also holds. In this example, notice that any even number of center runs with half assigned to

block 1 and the other half to block 2 results in orthogonal blocking. In the example, the factorial design points are assigned to one block and the axial points to the other block. In this case, condition 1 will hold automatically. The values for  $\alpha$  and  $n_c$  are chosen to achieve condition 2.

As a second example let us consider, again, a CCD but assume  $k=3$  and suppose that three blocks are required. The design is as follows:

| Block 1 |       |       | Block 2 |       |       | Block 3   |           |           |
|---------|-------|-------|---------|-------|-------|-----------|-----------|-----------|
| $x_1$   | $x_2$ | $x_3$ | $x_1$   | $x_2$ | $x_3$ | $x_1$     | $x_2$     | $x_3$     |
| -1      | -1    | -1    | 1       | -1    | -1    | $-\alpha$ | 0         | 0         |
| 1       | 1     | -1    | -1      | 1     | -1    | $\alpha$  | 0         | 0         |
| 1       | -1    | 1     | -1      | -1    | 1     | 0         | $-\alpha$ | 0         |
| -1      | 1     | 1     | 1       | 1     | 1     | 0         | $\alpha$  | 0         |
| 0       | 0     | 0     | 0       | 0     | 0     | 0         | 0         | $-\alpha$ |
| 0       | 0     | 0     | 0       | 0     | 0     | 0         | 0         | $\alpha$  |
|         |       |       |         |       |       | 0         | 0         | 0         |
|         |       |       |         |       |       | 0         | 0         | 0         |
|         |       |       |         |       |       | 0         | 0         | 0         |

Here we have divided the  $2^3 = 8$  factorial points into two blocks with  $I=ABC$  as the defining relation. The reader can again verify that condition 1 for orthogonal blocking holds. We can solve for  $\alpha$  that allows condition 2 to hold:

$$\frac{2\alpha^2}{9} = \frac{4}{6}$$

which result in  $\alpha = \sqrt{3}$ . As a result, we have a spherical CCD with  $\alpha = \sqrt{3}$  (the recommended choice for  $\alpha$  in Section 8.4.3), which blocks orthogonally in three blocks.

As the reader studies these two examples and the use of conditions 1 and 2, two principles obviously emerge: When one requires two blocks, there should be a factorial block and an axial block. In the case of three blocks, the factorial portion is divided into two blocks and the axial portion remains a single block. The partitioning into blocks of the factorial portion must result in a guaranteed orthogonality between blocks and the two factor interaction terms and, of course, orthogonality between blocks and linear coefficients. The numbers of blocks with the CCD are 2, 3, 5, 9, 17, ... . *There is never a subdivision of the axial block.* The axial points are always a single block. Table 8.8 provides more details of the blocking options for the CCD with different numbers of factors. The flexibility in the choice of center runs and the value of  $\alpha$  allows us to achieve orthogonal blocking or near-orthogonal blocking in many situations. At this point, we deal briefly in the special case of two blocks to give the reader further insight into the choice of  $\alpha$  and the assignment of center runs to blocks.

**Orthogonal Blocking in Two Blocks with the CCD** Suppose we denote by  $F_0$  the number of center runs in the factorial block, and by  $a_0$  the number of center runs in the axial

**TABLE 8.8 Design Details for Some Useful Central Composite Designs that Block Orthogonally**

| $k$   | 2  | 3  | 4  | 5  | 5 ( $\frac{1}{2}$ rep) | 6  | 6 ( $\frac{1}{2}$ rep) | 7   | 7 ( $\frac{1}{2}$ rep) |
|---|----|----|----|----|------------------------|----|------------------------|-----|------------------------|
| Total # of blocks                               | 2  | 3  | 3  | 5  | 2                      | 9  | 3                      | 17  | 9                      |
| Total # of points in design                     | 14 | 20 | 30 | 54 | 33                     | 90 | 54                     | 169 | 90                     |
| <i>Factorial block(s)</i>                       |    |    |    |    |                        |    |                        |     |                        |
| $F$ : # of points in factorial portion          | 4  | 8  | 16 | 32 | 16                     | 64 | 32                     | 128 | 64                     |
| # of blocks in factorial portion                | 1  | 2  | 2  | 4  | 1                      | 8  | 2                      | 16  | 8                      |
| # points in each block from factorial           | 4  | 4  | 8  | 8  | 16                     | 8  | 16                     | 8   | 8                      |
| $F_0$ : Number of added center points per block | 3  | 2  | 2  | 2  | 6                      | 1  | 4                      | 1   | 1                      |
| Total # of points per block                     | 7  | 6  | 10 | 10 | 22                     | 9  | 20                     | 9   | 9                      |
| <i>Axial block</i>                              |    |    |    |    |                        |    |                        |     |                        |
| $2k$ : # of axial points                        | 4  | 6  | 8  | 10 | 10                     | 12 | 12                     | 14  | 14                     |
| $\alpha_0$ : # of added center points           | 3  | 2  | 2  | 4  | 1                      | 6  | 2                      | 11  | 4                      |
| Total # of points in block                      | 7  | 8  | 10 | 14 | 11                     | 18 | 14                     | 25  | 18                     |

Source: G. E. P. Box and J. S. Hunter (1957).

block. Condition 1 holds when we assign the factorial and axial points as the two blocks. Condition 2 holds if

$$\frac{\sum_{\text{ax. bl.}} x_{iu}^2}{\sum_{\text{fac. bl.}} x_{iu}^2} = \frac{2k + a_0}{F + F_0}, \quad i = 1, 2, \dots, k \tag{8.30}$$

where, of course,  $F = 2^k$  or  $F = 2^{k-1}$  (resolution V fraction). The value for  $\alpha$  that results in orthogonal blocking is developed from Equation 8.30. From Equation 8.30 we have

$$\frac{2\alpha^2}{F} = \frac{2k + a_0}{F + F_0}$$

and the solution for  $\alpha$  is given by

$$\alpha = \sqrt{\frac{F(2k + a_0)}{2(F + F_0)}} \tag{8.31}$$

**General Recommendations for Blocking with the CCD** There are many instances in which the practitioner using RSM can achieve orthogonal blocking and still have designs that contain design parameters near those recommended in Section 8.4.3 when the region is spherical. We have already seen this to be the case with the  $k = 2$  and  $k = 3$  CCD. A value of  $\alpha = \sqrt{2}$  for the case of two blocks produces orthogonal blocking as long as center runs are evenly divided between the two blocks.

The case of the cuboidal region deserves special attention. In order to achieve orthogonal blocking for (say) a face-centered cube, one must resort to badly disproportionate block

sizes, which may often be a problem. For example, the following face-centered cube blocks orthogonally for  $k = 2$ :

$$\mathbf{D} = \begin{matrix} & x_1 & x_2 & & \\ & \left[ \begin{array}{cc} -1 & -1 \\ -1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \dots\dots\dots \\ -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{array} \right] & & \left. \begin{array}{l} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{array} \right\} & \begin{array}{l} \text{Block 1} \\ \\ \\ \\ \\ \\ \\ \\ \\ \text{Block 2} \end{array} \end{matrix}$$

For three variables, if one requires say three blocks, the two factorial blocks require eight center runs apiece if one requires that  $\alpha = 1.0$  in a cuboidal region. As a result, when ranges for the experiment seem to suggest the need for a cuboidal design region and orthogonal blocking is required, one might need to consider a spherical design in order that blocking be better accommodated.

In the case of spherical designs, many practical designs exist. Table 8.8 shows a working table of useful second-order designs that block orthogonally. For designs in which “ $\frac{1}{2}$  rep” is indicated, the factorial portion is a one-half fraction of the  $2^k$ . In all cases with multiple blocks from the factorial portion, defining contrasts are chosen so that three-factor interactions or higher are confounded with blocks.

**Blocking with the Box–Behnken Design** The flexibility of the CCD makes it an attractive choice when blocking is required in an RSM analysis. However, there are other second-order designs that block orthogonally. Included among these is the class of BBDs for  $k = 4$  and  $k = 5$ . Consider first the design matrix in Equation 8.16 for the  $k = 4$  BBD. Using  $\pm 1$  notation to indicate the  $2^2$  factorial structure among pairs of variables, we have

$$\mathbf{D} = \begin{matrix} & x_1 & x_2 & x_3 & x_4 & & \\ & \left[ \begin{array}{cccc} \pm 1 & \pm 1 & 0 & 0 \\ 0 & 0 & \pm 1 & \pm 1 \\ 0 & 0 & 0 & 0 \\ \dots\dots\dots \\ \pm 1 & 0 & 0 & \pm 1 \\ 0 & \pm 1 & \pm 1 & 0 \\ 0 & 0 & 0 & 0 \\ \dots\dots\dots \\ \pm 1 & 0 & \pm 1 & 0 \\ 0 & \pm 1 & 0 & \pm 1 \\ 0 & 0 & 0 & 0 \end{array} \right] & & \left. \begin{array}{l} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{array} \right\} & \begin{array}{l} \text{Block 1} \\ \\ \\ \\ \text{Block 2} \\ \\ \\ \text{Block 3} \end{array} \end{matrix}$$

The requirement on the vectors of center runs is that each block contain the same number of center runs. This, of course, results in equal block sizes, a requirement that results from the equal sum of squares for each factor in each block.

The  $k = 5$  BBD blocks orthogonally in two blocks. Consider the design of Equation 8.17. This design can be partitioned into two parts so that conditions for orthogonal blocking hold. Again, the center runs should be even in number, and half of them should be assigned to each block. The reader can easily verify that conditions 1 and 2 for orthogonal blocking both hold (see Exercise 8.27).

***Orthogonal Blocking with Definitive Screening Designs*** The flexibility of the Definitive Screening Designs to fit different models and to screen a large number of potential terms in a model has already been shown in Sections 8.3.4 and 8.4.8. Jones and Nachtsheim (2015) describe how orthogonal blocking of DSDs can be structured depending on the nature of the model under consideration.

If a first-order screening model is of primary importance, there are many potential DSDs with different block sizes that can be constructed. The basic structure of the DSD with the foldover pairs resulting from the form of the design based on the conference matrix means that as long as the pairs of runs with the same factor set to zero are in the same block, then the blocking effects will be orthogonal to the main effects. The inclusion of a center run in a block does not impact this orthogonality. This flexibility allows a screening design implementation of a DSD to be run to match almost any design constraints that the blocking may require.

When we consider first-order with interaction or second-order models, achieving orthogonal blocking becomes much more difficult, because of the large number of different terms that might be included in the final model based on common analysis strategies used for non-regular designs. Hence, the goal of orthogonal blocking with the DSDs might be weakened to still be able to achieve nonsingular estimation of the model terms. This would mean that although the terms are not orthogonal, or independent of the blocking effects, they are still estimable. For more details on options for blocking with DSDs, see Jones and Nachtsheim (2015).

***Form of Analysis with a Blocked Experiment*** The first and foremost topic of discussion in this area should be the model. It is important to understand that we are assuming that blocks enter the model as additive effects. This was outlined in the model form in Equation 8.25. It is assumed that there is no interaction between blocks and other model terms. This concept should be thoroughly discussed among the practitioners involved. If it is clear that blocks interact with the factors, then it means that the analysis should involve the fitting of a separate response surface for each block. This often results in difficulty regarding interpretation of results. Another important issue centers around what is truly meant by orthogonal blocking. As we indicated earlier, orthogonal means that block effects are orthogonal to model coefficients. However, the final fitted model will obviously include estimates of block effects. The fitted model is not at all unlike an analysis of covariance (ANCOVA) model where blocks play the role of treatments and one is assuming that for each block the model coefficients are the same (quite like assuming equal slopes in

ANCOVA). But, of course, the presence of block effects implies that we have multiple intercepts, one for each block. Thus, for, say, three blocks we have

$$\begin{aligned}\hat{y}_{\text{block}1} &= \hat{\beta}_{0,1} + f(\mathbf{x}) \\ \hat{y}_{\text{block}2} &= \hat{\beta}_{0,2} + f(\mathbf{x}) \\ \hat{y}_{\text{block}3} &= \hat{\beta}_{0,3} + f(\mathbf{x})\end{aligned}$$

where  $f(\mathbf{x})$  is a second-order function containing linear quadratic, and interaction terms, all of which are the same for each block. Now, the coefficients (apart from the intercepts) can be computed as if there were no blocking. As a result, the stationary point, canonical analysis, and ridge analysis path are computed by ignoring blocking, because none of these computations involves the intercepts. Then for the purpose of computing the predicted response, separate intercepts are merely found from the overall intercept and block effects. One must understand that the essence of the RSM analysis is intended to be a description of the system *in spite of* blocking; and blocks are assumed to have no effect on the nature and shape of the response surface. The following example should be beneficial.

**Example 8.7 The Analysis of a Blocked CCD** In a chemical process it is important to fit a response surface and find conditions on time ( $x_1$ ) and temperature ( $x_2$ ) that give a high yield. It is important that a yield of at least 80% be achieved. A list of natural and coded variables appears in Table 8.9. The design requires 14 experimental runs. However, two batches of raw materials must be used, so seven experimental runs were made for each batch. The experiment was run in two blocks, and the order of the runs was randomized in each block. The factorial runs plus three center runs were made in one block (batch), and the axial runs plus three center runs were made in the second block (batch). The resulting design does block orthogonally.

**TABLE 8.9 Natural and Design Levels for Chemical Reaction Experiment**

| Natural Variables |             | Design Units |        | Block | Yield Response |
|-------------------|-------------|--------------|--------|-------|----------------|
| Time              | Temperature | $x_1$        | $x_2$  |       |                |
| 80                | 170         | -1           | -1     | 1     | 80.5           |
| 80                | 180         | -1           | 1      | 1     | 81.5           |
| 90                | 170         | 1            | -1     | 1     | 82.0           |
| 90                | 180         | 1            | 1      | 1     | 83.5           |
| 85                | 175         | 0            | 0      | 1     | 83.9           |
| 85                | 175         | 0            | 0      | 1     | 84.3           |
| 85                | 175         | 0            | 0      | 1     | 84.0           |
| 85                | 175         | 0            | 0      | 2     | 79.7           |
| 85                | 175         | 0            | 0      | 2     | 79.8           |
| 85                | 175         | 0            | 0      | 2     | 79.5           |
| 92.070            | 175         | 1.414        | 0      | 2     | 78.4           |
| 77.930            | 175         | -1.414       | 0      | 2     | 75.6           |
| 85                | 182.07      | 0            | 1.414  | 2     | 78.5           |
| 85                | 167.93      | 0            | -1.414 | 2     | 77.0           |



**TABLE 8.10 Design-Expert Analysis of the Blocked CCD in Example 7.6 Using Design-Expert**

| Response                        | Yield          | <i>Sequential Model Sum of Squares</i> |                    |                     |          |  |           |
|---------------------------------|----------------|--|--------------------|---------------------|----------|--|-----------|
| Source                          | Sum of Squares | DF                                     | Mean Square        | F-Value             | Prob > F |  |           |
| Mean                            | 90916.80       | 1                                      | 90916.80           |                     |          |  |           |
| Block                           | 69.53          | 1                                      | 69.53              |                     |          |  |           |
| Linear                          | 9.63           | 2                                      | 4.81               | 2.67                | 0.1179   |  |           |
| 2FI                             | 0.063          | 1                                      | 0.063              | 0.031               | 0.8635   |  |           |
| Quadratic                       | 17.79          | 2                                      | 8.90               | 334.32              | <0.0001  |  | Suggested |
| Cubic                           | 0.044          | 2                                      | 0.022              | 0.78                | 0.5067   |  | Aliased   |
| Residual                        | 0.14           | 5                                      | 0.028              |                     |          |  |           |
| Total                           | 91014.00       | 14                                     | 6501.00            |                     |          |  |           |
| <i>Lack of Fit Tests</i>        |                |  |                    |                     |          |  |           |
| Linear                          | 17.91          | 6                                      | 2.98               | 89.53               | 0.0003   |  |           |
| 2FI                             | 17.84          | 5                                      | 3.57               | 107.07              | 0.0002   |  |           |
| Quadratic                       | 0.053          | 3                                      | 0.018              | 0.53                | 0.6859   |  | Suggested |
| Cubic                           | 8.571E-003     | 1                                      | 8.571E-003         | 0.26                | 0.6388   |  | Aliased   |
| Pure Error                      | 0.13           | 2                                      | 0.033              |                     |          |  |           |
| <i>Model Summary Statistics</i> |                |  |                    |                     |          |  |           |
| Source                          | Std. Dev.      | R-Squared                              | Adjusted R-Squared | Predicted R-Squared | PRESS    |  |           |
| Linear                          | 1.34           | 0.3479                                 | 0.2175             | -0.2812             | 35.44    |  |           |
| 2FI                             | 1.41           | 0.3502                                 | 0.1336             | -0.9332             | 53.48    |  |           |
| Quadratic                       | 0.16           | 0.9933                                 | 0.9885             | 0.9726              | 0.76     |  | Suggested |
| Cubic                           | 0.17           | 0.9949                                 | 0.9877             | 0.9451              | 1.52     |  | Aliased   |

(continued)

TABLE 8.10 Continued

| Response   |                      | Yield          |                |             |                |          |
|--|----------------------|----------------|----------------|-------------|----------------|----------|
| ANOVA for Response Surface Quadratic Model                 |                      | Sum of Squares | DF             | Mean Square | F-Value        | Prob > F |
| <i>Analysis of Variance Table [Partial Sum of Squares]</i> |                      |                |                |             |                |          |
| Source   |                      |                |                |             |                |          |
| Block  |                      | 69.53          | 1              | 69.53       |                |          |
| Model  |                      | 27.48          | 5              | 5.50        | 206.55         | <0.0001  |
| A  |                      | 6.96           | 1              | 6.96        | 261.43         | <0.0001  |
| B  |                      | 2.67           | 1              | 2.67        | 100.33         | <0.0001  |
| A <sup>2</sup>   |                      | 12.64          | 1              | 12.64       | 475.06         | <0.0001  |
| B <sup>2</sup>   |                      | 6.43           | 1              | 6.43        | 241.76         | <0.0001  |
| AB   |                      | 0.063          | 1              | 0.063       | 2.35           | 0.1692   |
| Residual   |                      | 0.19           | 7              | 0.027       |                |          |
| Lack of Fit  |                      | 0.053          | 3              | 0.018       | 0.53           | 0.6859   |
| Pure Error   |                      | 0.13           | 4              | 0.033       |                |          |
| Cor total  |                      | 97.20          | 13             |             |                |          |
| Std. Dev.  |                      | 0.16           |                |             |                |          |
| Mean   |                      | 80.59          |                |             |                |          |
| C.V.   |                      | 0.20           |                |             |                |          |
| PRESS  |                      | 0.76           |                |             |                |          |
|  |                      |                |                |             | R-Squared      | 0.9933   |
|  |                      |                |                |             | Adj R-Squared  | 0.9885   |
|  |                      |                |                |             | Pred R-Squared | 0.9726   |
|  |                      |                |                |             | Adeq Precision | 72.761   |
| Factor   | Coefficient Estimate | DF             | Standard Error | 95% CI      |                | VIF      |
| Intercept  | 81.87                | 1              | 0.067          | Low         | High           |          |
| Block 1  | 2.23                 | 1              |                | 81.71       | 82.02          |          |
| Block 2  | -2.23                |                |                |             |                |          |
| A-Temp   | 0.93                 | 1              | 0.058          | 0.80        | 1.07           | 1.00     |

|                |       |   |       |        |       |      |
|----------------|-------|---|-------|--------|-------|------|
| B-Temp         | 0.58  | 1 | 0.058 | 0.44   | 0.71  | 1.00 |
| A <sup>2</sup> | -1.31 | 1 | 0.060 | -1.45  | -1.17 | 1.01 |
| B <sup>2</sup> | -0.93 | 1 | 0.060 | -1.08  | -0.79 | 1.01 |
| AB             | 0.13  | 1 | 0.082 | -0.068 | 0.32  | 1.00 |

**Final Equation in Terms of Coded Factors:**

|        |                  |
|--------|------------------|
| Yield  | =                |
| +81.87 |                  |
| +0.93  | * A              |
| +0.58  | * B              |
| -1.31  | * A <sup>2</sup> |
| -0.93  | * B <sup>2</sup> |
| +0.13  | * A*B            |

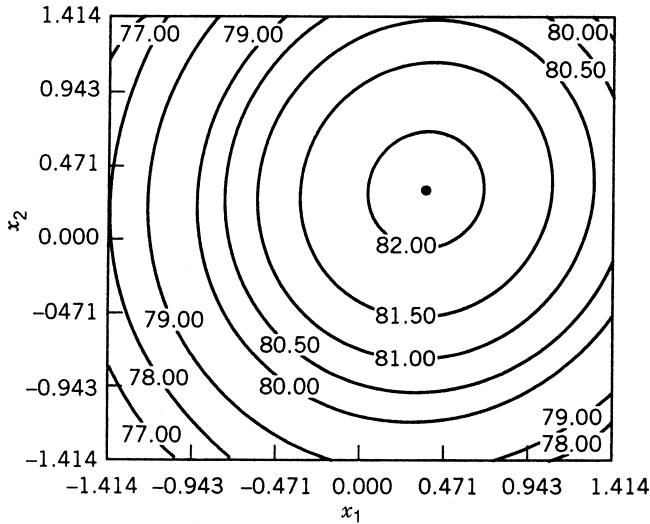
**Final Equation in Terms of Actual Factors:**

|                |                     |
|----------------|---------------------|
| Yield          | =                   |
| -1401.27035    |                     |
| +8.20816       | * Time              |
| +12.75720      | * Temp              |
| -0.052333      | * Time <sup>2</sup> |
| -0.037333      | * Temp <sup>2</sup> |
| +5.000000E-003 | * Time*Temp         |

“*Sequential Model Sum of Squares*”: Select the highest order polynomial where the additional terms are significant.

“*Lack of Fit Test*”: Want the selected model to have insignificant lack-of-fit.

“*Model Summary Statistics*”: Focus on the model maximizing the “Adjusted R-Squared” and the “Predicted R-Squared.”



**Figure 8.21** Contour plot of yield for Example 8.7.

A second-order analysis was performed. Table 8.10 shows the analysis produced by Design-Expert, and Fig. 8.21 is a contour plot of the fitted model shown in the coded units. The stationary point is a maximum. The fitted model is

$$\hat{y} = 81.87 + 0.93x_1 + 0.58x_2 - 1.31x_1^2 - 0.93x_2^2 + 0.13x_1x_2$$

Note that the  $P$ -values indicate that one model term ( $x_1x_2$ ) is insignificant. The data were analyzed *as if no blocking had been done*. This renders coefficients (and canonical analysis, if one is performed) correct, but standard errors are inflated because of variability due to blocks. The appropriate location of the maximum response is given by

$$\begin{aligned}x_{1,s} &= 0.372 \\x_{2,s} &= 0.335\end{aligned}$$

in our design units. The predicted value at the stationary point is 82.136.

Further examination of the Design-Expert output can determine the block effects. Note that while the overall intercept is 81.87, the estimates of block effects are 2.23 and  $-2.23$  for block 1 and block 2, respectively. As a result, the two separate intercepts are  $81.867 \pm 2.23 = 84.097$  and  $79.637$ , respectively.

## EXERCISES

**8.1** Consider the design of an experiment to fit the first-order response model

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \beta_5x_5 + \varepsilon$$

The design used is a  $2^{5-2}$  fractional factorial with defining relations

$$I = ABC = x_1x_2x_3$$

$$I = CDE = x_3x_4x_5$$

The factors are quantitative, and thus we use the notation  $x_1, x_2, x_3, \dots$  rather than  $A, B, C, \dots$

- (a) Construct the matrix  $\mathbf{X}$  for the first-order model using the design described above.
- (b) Is the design orthogonal? Explain why or why not.
- (c) Is the design a variance-optimal design—that is, one that results in minimum values of  $N \text{Var}(b_i)/\sigma^2$ . Note that the weight  $N$  is used in order to take sample size into account.
- 8.2** Consider the situation of Exercise 8.1. Suppose we use four center runs to augment the  $2^{5-2}$  design constructed.
- (a) What is the advantage of using the center runs?
- (b) Will the resulting 12-run design be orthogonal? Explain.
- (c) Will the resulting design be variance-optimal?
- (d) If your answer to (c) is no, give a 12-run first-order design with values of  $\text{Var}(b_i)/\sigma^2$  smaller than the 12-run design with center runs.
- (e) Give the variances of regression coefficients for both designs—that is, the  $2^{5-2}_{\text{III}}$  with four center runs and your design in (d).
- 8.3** Consider again the design constructed in Exercise 8.1. Suppose the fitted model is first-order and the analysis reveals significant lack of fit and a large effect for the interaction  $x_1x_4 = x_2x_5$ . As a result, a second phase would suggest an augmentation of the design. Suppose the second phase is a fold-over of the design in Exercise 8.1. Is the resulting design orthogonal and thus variance-optimal for the model

$$y_u = \beta_0 + \sum_{i=1}^5 \beta_i x_{ui} + \sum_{i < j=2}^5 \beta_{ij} x_{ui} x_{uj} + \varepsilon_u,$$

$$u = 1, 2, 3, \dots, 16$$

Explain why or why not.

- 8.4** Consider the  $L_{27}$  design (in coded form) discussed in Exercise 6.12. Suppose we consider this design for fitting the first-order model

$$y_u = \beta_0 + \sum_{i=1}^3 \beta_i x_{ui} + \varepsilon_u, \quad u = 1, 2, \dots, 27$$

- (a) Is the design first-order orthogonal? Explain.
- (b) Is the design variance-optimal? Explain.
- (c) Compare the  $N \text{Var}(b_i)/\sigma^2$  using the  $L_{27}$  with a  $2^3$ . Compare with a Plackett–Burman 12-run design. Compare with a Plackett–Burman 28-run design.

- (d) Your answer to (c) above should be that the Plackett–Burman 12-run design and  $2^3$  design for a first-order model are *variance-equivalent on a per observation basis*. What are extra advantages enjoyed by the  $2^3$  over the Plackett–Burman 12-run design?
- 8.5 Consider a situation involving seven design variables when, in fact, a first-order model is required but only eight runs can be used. Design a saturated eight-run design that is variance-optimal.
- 8.6 Consider a situation in which five factors are to be studied and a first-order model is postulated though it is not quite clear that this is the correct model. Sixteen design runs are to be used. The designs to be considered are as follows:
- (a)  $2_{III}^{5-2}$  replicated
- (b)  $2_V^{5-1}$
- (c) A Plackett–Burman 12-run design augmented with four center runs.
- Discuss the advantages and disadvantages of the three designs. Use the terms *variance-optimal* and *model misspecification* in your discussion.
- 8.7 Consider the “model inadequacy” material in Section 8.2.2. The material can be used to nicely illustrate the aliasing ideas discussed in Chapter 4. Aliasing plays an important role when one deals with fractional factorials for fitting first-order and first-order-plus-interaction response surface models. Suppose one fits a first-order model using a  $2_{III}^{3-1}$  with defining relation

$$I = x_1x_2x_3$$

However, suppose the true model is

$$E(y) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{12}x_1x_2 + \beta_{13}x_1x_3 + \beta_{23}x_2x_3$$

Using Equation 8.2, we obtain

$$\mathbf{X}_1 = \begin{matrix} & x_1 & x_2 & x_3 \\ \begin{bmatrix} 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}, & \mathbf{X}_2 = \begin{matrix} x_1x_2 & x_1x_3 & x_2x_3 \\ \begin{bmatrix} -1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \end{matrix}$$

$$\beta_1 = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}, \quad \beta_2 = \begin{bmatrix} \beta_{12} \\ \beta_{13} \\ \beta_{23} \end{bmatrix}$$

Use the expression

$$\begin{aligned} E(\mathbf{b}_1) &= \beta_1 + (\mathbf{X}'_1\mathbf{X})^{-1}\mathbf{X}'_1\mathbf{X}_2\mathbf{y} \\ &= \beta_1 + \mathbf{A}\mathbf{y} \end{aligned}$$

to verify the following expected values:

$$\begin{aligned} E(b_0) &= \beta_0 \\ E(b_1) &= \beta_1 - \beta_{23} \\ E(b_2) &= \beta_2 - \beta_{13} \\ E(b_3) &= \beta_3 - \beta_{12} \end{aligned}$$

**8.8** In Exercise 8.7, if the true model is given by

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3$$

show that the expected values are given by

$$\begin{aligned} E(b_0) &= \beta_0 - \beta_{123} \\ E(b_1), E(b_2), E(b_3) &\text{ as in Exercise 8.7} \end{aligned}$$

Exercises 8.7 and 8.8 illustrate the effect of aliasing in the sense of a regression model. Thus, in the case of these two exercises, the aliasing of  $x_1$  with  $x_2 x_3$  implies that the coefficient  $b_1$  is not an unbiased estimator of  $\beta_1$  but is biased by  $\beta_{23}$ , the coefficient of  $x_2 x_3$ . Indeed, the coefficient of the interaction described by the defining contrast, namely  $\beta_{123}$ , biases the estimated intercept  $\beta_0$ .

**8.9** Consider the same situation as in Exercise 8.7. However, the *other fraction* given by  $I = x_1 x_2 x_3$  is used. If the fitted first-order model is incorrect and the true model is as stated in Exercise 8.8, show that the biases in the coefficients are indicated by

$$\begin{aligned} E(b_0) &= \beta_0 + \beta_{123} \\ E(b_1) &= \beta_1 + \beta_{23} \\ E(b_2) &= \beta_2 + \beta_{13} \\ E(b_3) &= \beta_3 + \beta_{12} \end{aligned}$$

**8.10** Consider a  $2^{5-2}$  fractional factorial with defining relations

$$\begin{aligned} I &= x_1 x_2 x_3 \\ I &= x_3 x_4 x_5 \end{aligned}$$

Suppose a first-order model is fitted to the data and the true model includes, additionally, all two-factor interactions. Use the expression

$$E(\mathbf{b}_1) = \beta_1 + \mathbf{A}\beta_2$$

to develop,  $E(b_1)$ ,  $E(b_2)$ ,  $E(b_3)$ ,  $E(b_4)$ , and  $E(b_5)$  where the fitted first-order model is

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + b_5 x_5$$

- 8.11** Consider Example 4.2 in Chapter 4. The strong effects are those for the factors  $A$ ,  $B$ ,  $C$ , and  $AB$ , and the fitted regression is given by

$$\hat{y} = 60.65625 + 11.60625x_1 + 34.03125x_2 + 10.45625x_3 + 6.58125x_1x_2$$

- (a) Verify the above coefficients by considering this five-parameter model in the form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

and by using

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

- (b) Is the design in Table 4.5 orthogonal for the model that was fitted? If so, verify it. Is the design variance-optimal?
- (c) Is the design orthogonal for a model containing  $x_1$ ,  $x_2$ ,  $x_3$ ,  $x_4$ ,  $x_5$ ,  $x_1x_2$ ,  $x_1x_3$ ,  $x_1x_4$ ,  $x_1x_5$ ,  $x_2x_3$ ,  $x_2x_4$ ,  $x_2x_5$ ,  $x_3x_4$ ,  $x_3x_5$ , and  $x_4x_5$ ?
- 8.12** Consider the design listed in Table 4.11 with the injection molding data. Answer the following questions.
- (a) For the model that is fitted, namely

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_{12}x_{12}$$

give the variance of each coefficient, assuming common error variance  $\sigma^2$ .

- (b) Using the error mean square from Table 4.12, give the estimated standard errors of all four coefficients.
- (c) In light of the plot in Figure 4.14, why are the results in (b) somewhat precarious?
- (d) Is this design orthogonal for a model containing all linear terms and the two-factor interactions  $x_1x_2$ ,  $x_1x_3$ ,  $x_1x_4$ ,  $x_1x_5$ , and  $x_1x_6$ ? If so, justify it; if not, explain why not.
- 8.13** In Section 4.5 of Chapter 4 the  $2_{III}^{7-4}$  design is discussed. Here we may be interested in building a regression model for studying seven factors. The design is given in Table 4.15. Suppose one fits a first-order model but, in fact, all two factor interactions are in the true model. Write out  $E(b_1)$ ,  $E(b_2)$ ,  $E(b_3)$ ,  $E(b_4)$ ,  $E(b_5)$ ,  $E(b_6)$ , and  $E(b_7)$ . Explain all your terms.
- 8.14** Consider Exercise 4.7 in Chapter 4.
- (a) Is the design first-order orthogonal?
- (b) From your answer in Exercise 4.7b, write a response surface model.
- (c) Give estimated standard errors of all coefficients.
- (d) From your model in Exercise 8.14b estimate the standard error of prediction at all design locations.
- (e) Compute the estimated standard error of prediction at the design center.



- 8.15** Consider Exercise 4.8 in which four center runs were added to the design in Exercise 4.7. Answer the questions asked in Exercise 8.14d and 8.14e again. Use the replication error variability (3 df) in computing your mean square error.
- 8.16** (a) Consider Exercise 4.11 in Chapter 4. For the data in Exercise 4.11a develop a first-order regression model. Estimate the standard error of prediction at each design point.  
 (b) Do all the work in (a) again after the second fraction in Exercise 4.11b is added. Comment on what was gained with the addition of the second fraction.
- 8.17** Suppose we are interested in fitting a response surface model in five design variables. The analyst is quite sure that the form of the model should be

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} \\ + \beta_{12} x_{i1} x_{i2} + \beta_{13} x_{i1} x_{i3} + \varepsilon_i$$

The analyst can use 12 experimental runs in the design. Discuss and compare the following candidate designs. Use the variances of individual coefficients in your discussion. However, do not let your discussion be confined to this criterion. The candidate designs are as follows:

$$\mathbf{D}_1 : 2_{III}^{5-2} \text{ with four center runs} \\ \mathbf{D}_2 : \text{Plackett–Burman with 12 runs}$$

- 8.18** Equation 8.11 is a very important expression. It gives the scaled prediction variance for a first-order orthogonal design in which all design points are at the  $\pm 1$  extremes. It also implies that this class of designs give equal prediction variance at any two locations that are on the same sphere. This implies that this class of designs is rotatable. Does the same property hold for orthogonal designs used to accommodate a model containing interaction? Illustrate with a  $2^2$  factorial and the model

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{12} x_{i1} x_{i2}$$

Compute  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  at the following locations on the same sphere:

- (i) 1, 1
- (ii) -1, 1
- (iii) 1, -1
- (iv) 1, 1
- (v)  $\sqrt{2}$ , 0
- (vi)  $-\sqrt{2}$ , 0
- (vii) 0,  $-\sqrt{2}$
- (viii) 0,  $\sqrt{2}$

Comment on your results.

- 8.19 A first-order orthogonal design with all points at  $\pm 1$  extremes remains rotatable for a first-order model if all points are replicated the same number of times. True or false? Justify your answer.
- 8.20 A second-order rotatable design with  $\alpha = \sqrt[4]{F}$  remains rotatable for the second-order model if all factorial and axial runs are replicated the same number of times. True or false? Justify your answer.
- 8.21 Consider the model in Exercise 8.17 along with designs  $D_1$  and  $D_2$ . Which design is superior using the scaled prediction variance criterion? Use as illustrations the following locations:

- (1, 1, 1, 1, 1)
- (-1, -1, -1, -1, -1)
- (1, 1, -1, -1, -1)
- (1, -1, -1, 1, 1)
- (1, -1, -1, -1, -1)
- (-1, 1, 1, 1, 1)

- 8.22 Consider a central composite design with  $\alpha = \sqrt{F}$ . Use the result in Equation 7.14 to show that the number of center runs has no effect on the rotatability property.
- 8.23 It is stated in Section 8.4.9 that for second-order models, orthogonal blocking is much more easily accommodated with the use of the spherical CCD than with use of the face-centered cube (cuboidal CCD). However, orthogonal blocking is accommodated with the face-centered cube when *axial points are replicated*. Show that the following design blocks orthogonally in two blocks:

$$\mathbf{D}_1 = \begin{matrix} & x_1 & x_2 & & \\ & \left[ \begin{array}{cc} -1 & -1 \\ -1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{array} \right] & & \left. \vphantom{\begin{array}{cc} -1 & -1 \\ -1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{array}} \right\} \text{Block 1} \\ & \dots\dots\dots & & \\ & \left[ \begin{array}{cc} -1 & 0 \\ -1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & -1 \\ 0 & 1 \\ 0 & 1 \end{array} \right] & & \left. \vphantom{\begin{array}{cc} -1 & 0 \\ -1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & -1 \\ 0 & 1 \\ 0 & 1 \end{array}} \right\} \text{Block 2} \end{matrix}$$

- 8.24** Consider Table 8.5. The designs suggested in the table are those that block orthogonally and are rotatable or near-rotatable. For example, consider the design under the column headed  $k = 5(\frac{1}{2}\text{rep})$ . Construct the design completely, and verify that it meets both conditions for orthogonal blocking.
- 8.25** Answer the same question as in Exercise 8.23, but use the  $k = 5$  line with four blocks from the factorial portion and one block from the axial portion.
- 8.26** The designs given in Table 8.8 are not the only ones that block orthogonally and are practical. The values for  $\alpha$  and  $n_c$  can be altered in a very flexible manner. Suppose, for example, that for  $k = 4$  the run size does not allow a design that is shown in the table. Rather, we must use

$$\mathbf{D} = \begin{array}{cccc}
 & x_1 & x_2 & x_3 & x_4 \\
 \left[ \begin{array}{cccc}
 -1 & -1 & -1 & -1 \\
 1 & 1 & -1 & -1 \\
 1 & -1 & 1 & -1 \\
 1 & -1 & -1 & 1 \\
 -1 & 1 & 1 & -1 \\
 -1 & 1 & -1 & 1 \\
 -1 & -1 & 1 & 1 \\
 1 & 1 & 1 & 1 \\
 0 & 0 & 0 & 0 \\
 \dots\dots\dots & & & & \\
 1 & -1 & -1 & -1 \\
 -1 & 1 & -1 & -1 \\
 -1 & -1 & 1 & -1 \\
 -1 & -1 & -1 & 1 \\
 1 & 1 & 1 & -1 \\
 1 & 1 & -1 & 1 \\
 1 & -1 & 1 & 1 \\
 -1 & 1 & 1 & 1 \\
 0 & 0 & 0 & 0 \\
 \dots\dots\dots & & & & \\
 -\alpha & 0 & 0 & 0 \\
 \alpha & 0 & 0 & 0 \\
 0 & -\alpha & 0 & 0 \\
 0 & \alpha & 0 & 0 \\
 0 & 0 & -\alpha & 0 \\
 0 & 0 & \alpha & 0 \\
 0 & 0 & 0 & -\alpha \\
 0 & 0 & 0 & \alpha
 \end{array} \right.
 \end{array}$$

This design can be made to block orthogonally. What value of  $\alpha$  results in orthogonal blocking?

- 8.27** Construct a Box–Behnken design in five variables that blocks orthogonally for a second-order model. Show the assignment of design points to blocks.

- 8.28** Compare a  $2^{6-3}$  fractional factorial design ( $I = ABD = ACE = BCF$ ) with four center runs (total of 12 runs), or a Plackett–Burman design with 12 runs for estimating the model

$$y = \beta_0 + \beta_1A + \beta_2B + \beta_3C + \beta_4D + \beta_5E + \beta_6F + \beta_{12}AB$$

Which is better for

- (a) estimating effects independently?
- (b) smallest variance for coefficient estimates?
- (c) testing lack of fit?
- (d) testing curvature?
- (e) estimating pure error?

Based on (a)–(e), which design do you prefer? Explain.

- 8.29** We are interested in studying the main effects ( $A, B, C, D, E$ ) and some two-way interactions ( $AB, AC$ ) for five factors by running a half fraction of a fractional factorial. We also have the restriction that we can only have eight observations per block.
- (a) Suggest a suitable design that will allow all effects to be estimated independently of each other. Give the defining equation, as well as what effect would be aliased with the block.
  - (b) List the combinations of  $A, B, C, D, E$  would be run in each of the blocks.
  - (c) If we assume a model with terms for  $A, B, C, D, E, AB, AC$ , and *Block*, what would the ANOVA table for this design look like? (Source, d.f.)
  - (d) Give the equation (using combinations of  $A, B, C, D, E$ ) for how you would calculate the effect for the  $AB$  interaction.
  - (e) If the true model had terms  $A, B, C, D, E, AB, AC$ , and *Block* as well as  $BC$  and  $CD$ , would your model be adequate to estimate all of the effects? Explain.
- 8.30** For a full  $2^3$  factorial with two center runs, we assume a first-order model with all two-way interactions is the correct model. If the correct model is actually, the first-order model with two-way interactions and a quadratic effect for factor  $A$ , find which estimates are unbiased?
- 8.31** For the following designs and specified models, give the appropriate design matrix, and show whether or not they are rotatable (work with design moments where convenient).
- (a) Plackett–Burman design (12 runs) for a first-order model for effects  $A, B, C, D, E$
  - (b) Factorial design for a first-order model with interactions for effects  $A, B, C$
  - (c) CCD design with half-fraction for the factorial portion,  $\alpha = 2$  and  $n_c = 3$  for a second-order model for effects  $A, B, C, D, E$
  - (d) BBD design with  $n_c = 3$  for a second-order model for effects  $A, B, C, D, E$
  - (e) An equiradial design with seven points for a second-order model for effects  $A$  and  $B$ .

- 8.32** Assume a second-order model with three factors ( $A, B, C$ ) is the model of interest. For the following designs, calculate the scaled predicted variance,  $SPV = (N\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2)$  along the lines  $(a, 0, 0)$ ,  $(a, a, 0)$ ,  $(a, -a, 0)$ ,  $(a, a, a)$ , and  $(a, -a, a)$  for different values of  $a$ . Select an appropriate range to consider and plot the results on a graph of distance from center versus  $SPV$ . Comment on your results, referring to rotability and desirability of the design.
- (a) CCD with  $\alpha = 1$ ,  $n_c = 2$
  - (b) CCD with  $\alpha = 1$ ,  $n_c = 4$
  - (c) BBD with  $n_c = 2$
  - (d) BBD with  $n_c = 4$

- 8.33** For a  $2^{4-1}$  fractional factorial design with defining equation  $I = ABCD$  and assumed first-order model with interactions  $AB, AC$ , and  $BC$ :
- (a) What is the “alias matrix” if a better model for the relationship between factors and response has all of the above terms as well as the interactions  $AD$  and  $ABC$ ? What coefficient estimates will not be unbiased?
  - (b) Now assume that the run order of the experiment was as follows:

$$ab, cd, ad, ac, (1), abcd, bc, bd.$$

If (unknown to the experimenter) conditions changed after the first four runs (treat this as two “self-imposed” blocks), what will the “alias matrix” look like (assume the initial model is otherwise correct)? What coefficient estimates will be unbiased?

- 8.34** An experimenter wants to run a screening experiment with six factors of interest and 16 runs. She *thinks* a first-order model is adequate for any effects that are present. Consider the following four designs:
- D**<sub>1</sub>: A  $2^{6-2}_{IV}$  fractional factorial.
  - D**<sub>2</sub>: A  $2^{6-3}_{III}$  fractional factorial two reps at each location.
  - D**<sub>3</sub>: A Plackett–Burman design with  $n = 12$  and four center runs.
  - D**<sub>4</sub>: A Plackett–Burman design with  $n = 12$  and four reps chosen at random locations.

Which is better for

- (a) estimating effects independently?
- (b) smallest variance for coefficient estimates?
- (c) testing lack of fit?
- (d) testing curvature?
- (e) estimating pure error?

Based on (a)–(e), which design do you prefer? Explain.

- 8.35** For an experiment with four factors ( $A, B, C, D$ ) in a spherical region, a CCD with  $\alpha = 2$  and  $n_c = 3$  is selected. It is assumed that a second-order model is adequate. Find the “alias matrix” if a better model for the relationship between factors and response includes the cubic terms  $A^3, B^3, C^3, D^3$ . Which terms in the second-order model will be unbiased?

- 8.36** For an experiment with four factors ( $A, B, C, D$ ) in a spherical region, a CCD with  $\alpha = 2$  and  $n_c = 3$  is selected. It is assumed that a second-order model is adequate. Find the “alias matrix” if a better model for the relationship between factors and response includes the third-order terms  $ABC, ABD, ACD, BCD$ . Which terms in the second-order model will be unbiased?
- 8.37** An experiment involving three factors in a spherical region can be run in blocks as large as eight runs. Three competing designs with orthogonal blocking are being considered:
- D<sub>1</sub>**: A three block design given in Table 8.8 with two factorial blocks of size 6 and the axial block of size 8.
- D<sub>2</sub>**: A three block design with two factorial blocks and an axial block, all of size 8.
- D<sub>3</sub>**: A two block design with one factorial block (of size 8) and an axial block of size 8.
- (a) For design **D<sub>2</sub>**, find the axial distance that makes the blocks orthogonal.
- (b) For design **D<sub>3</sub>**, find the axial distance that makes the blocks orthogonal.
- (c) Based on a comparison of estimation of pure error, location of the axial runs, and overall size of the designs, rank the different designs.
- (d) Which design would you recommend for the experiment?
- 8.38** An experiment involving four factors in a spherical region can be run in blocks as large as 12 runs. Three competing designs with orthogonal blocking are being considered:
- D<sub>1</sub>**: A three-block design given in Table 8.8 with two factorial blocks of size 10 and the axial block of size 10.
- D<sub>2</sub>**: A three-block design with two factorial blocks and an axial block, all of size 12.
- D<sub>3</sub>**: A three-block design with two factorial blocks of size 12 and an axial block of size 10.
- (a) For design **D<sub>2</sub>**, find the axial distance that makes the blocks orthogonal.
- (b) For design **D<sub>3</sub>**, find the axial distance that makes the blocks orthogonal.
- (c) Based on a comparison of estimation of pure error, location of the axial runs, and overall size of the designs, rank the different designs.
- (d) Which design would you recommend for the experiment?
- 8.39** Maruyama et al. (2013) conducted a study on the impact of replacing some portion of wheat flour in bread with other ingredients to determine the impact on several properties of the bread. Data from the experiment is shown in Table E8.1, with the true factor levels for the three factors (Chia seeds, Carrot leaves and Baking time) shown in parentheses.
- (a) What type of design was used for the experiment?
- (b) What is the scaled prediction variance for the design, for the following design locations:
- (i) Chia seed = 1.5 g, Carrot leaves = 1.5 g, Baking time = 70 min
- (ii) Chia seed = 1.0 g, Carrot leaves = 2.0 g, Baking time = 80 min
- (iii) Chia seed = 1.5 g, Carrot leaves = 1.0 g, Baking time = 60 min

**TABLE E8.1** Data for Exercises 8.39, 8.40, and 8.41

| Run | Chia<br>Seeds (g) | Carrot<br>leaves (g) | Baking<br>Time (min) | Moisture | Ash  | Crude<br>Protein | Total<br>Lipid |
|-----|-------------------|----------------------|----------------------|----------|------|------------------|----------------|
| 1   | -1 (1.0)          | -1 (1.0)             | 0 (70)               | 36.76    | 1.66 | 7.4              | 2.88           |
| 2   | 1 (2.0)           | -1 (1.0)             | 0 (70)               | 34.07    | 1.79 | 7.63             | 2.98           |
| 3   | -1 (1.0)          | 1 (2.0)              | 0 (70)               | 35.15    | 1.73 | 7.48             | 3.5            |
| 4   | 1 (2.0)           | 1 (2.0)              | 0 (70)               | 35.76    | 1.99 | 8.12             | 3.97           |
| 5   | -1 (1.0)          | 0 (1.5)              | -1 (60)              | 36.16    | 1.73 | 7.44             | 3.16           |
| 6   | 1 (2.0)           | 0 (1.5)              | -1 (60)              | 37.28    | 1.86 | 7.86             | 3.29           |
| 7   | -1 (1.0)          | 0 (1.5)              | 1 (80)               | 33       | 1.69 | 7.43             | 3.19           |
| 8   | 1 (2.0)           | 0 (1.5)              | 1 (80)               | 33.43    | 1.94 | 7.85             | 3.6            |
| 9   | 0 (1.5)           | -1 (1.0)             | -1 (60)              | 35.36    | 1.68 | 7.48             | 2.91           |
| 10  | 0 (1.5)           | 1 (2.0)              | -1 (60)              | 35.66    | 1.89 | 7.76             | 3.05           |
| 11  | 0 (1.5)           | -1 (1.0)             | 1 (80)               | 35.22    | 1.73 | 7.47             | 3.02           |
| 12  | 0 (1.5)           | 1 (2.0)              | 1 (80)               | 34.36    | 1.9  | 7.73             | 3.24           |
| 13  | 0 (1.5)           | 0 (1.5)              | 0 (70)               | 33.89    | 1.85 | 7.6              | 3.35           |
| 14  | 0 (1.5)           | 0 (1.5)              | 0 (70)               | 33.2     | 1.84 | 7.61             | 3.31           |
| 15  | 0 (1.5)           | 0 (1.5)              | 0 (70)               | 33.57    | 1.84 | 7.6              | 3.28           |

- (c) Do you have concerns about predicting the responses at any of the design locations in (b)? Why or why not?
- (d) Suppose that the experimenter had wanted to use a Central Composite design instead. Using the same ranges of the factors and same-sized experiment as in Table E8.1, create a face-centered CCD.
- (e) Compare the scaled prediction variance values for the design locations in (b). Which design do you prefer? Why?
- (f) Create a Fraction of Design Space plot for each design. Which design do you prefer? Why?
- 8.40** Reconsider the experiment in Exercise 8.39 and Table E8.1. The primary characteristic of interest in the experiment is the amount of crude protein.
- (a) Fit a second-order model for the response crude protein.
- (b) At which location in the design space is the estimated amount of crude protein maximized? Think carefully about how you define the design space.
- (c) Find a 95% confidence region for the amount of crude protein at the location found in (b).
- (d) Find all of the locations in the design space for which their 95% confidence intervals include the point estimate for the maximum amount of crude protein found in (b).
- (e) Find all of the locations in the design space for which their 95% prediction intervals for a new observation include the point estimate for the maximum amount of crude protein found in (b).
- 8.41** Reconsider the experiment in Exercise 8.40 and Table E8.1. While the primary goal of the optimization is to maximize the amount of crude protein, secondary objectives include having sufficient moisture and minimizing the amount of ash.

- (a) Fit second-order response surface models to both moisture and ash.
- (b) Find the region in the design space where moisture is estimated to have a value of at least 35.
- (c) Find the location in the design space that minimizes the amount of ash.
- (d) Based on the results of (b) and (c), comment on the amount of trade-off between maximizing crude protein (from Exercise 8.40), maximizing moisture and minimizing ash.
- (e) Find the location in the design space that maximizes crude protein, where moisture has a value of at least 35.
- (f) Find the Pareto front for locations in the design space that balance minimizing ash and maximizing crude protein.

**8.42** Amenaghawon et al. (2013) consider optimizing the production of citric acid with a Box–Behnken design involving three factors. They consider carbon substrate (*Aspergillus niger*) concentration (X1), broth pH (X2), and fermentation temperature (X3) as the three factors. The ranges for each of the factors are as follows:

|                                 |           |
|---------------------------------|-----------|
| Carbon substrate concentration: | 30–50 g/L |
| Broth pH:                       | 2–5       |
| Fermentation temperature:       | 25°–35°C  |

- (a) Fit a second order model for citric acid based on the three factors.
- (b) If the goal of the experiment is to maximize the amount of citric acid produced, what is the optimal combination of carbon substrate concentration, broth pH and fermentation temperature within the design region? What is the value of the response at this location?
- (c) If it were possible to expand the range of each of the factors to

|                                 |           |
|---------------------------------|-----------|
| Carbon substrate concentration: | 25–55 g/L |
| Broth pH:                       | 1–6       |
| Fermentation temperature:       | 20°–40°C  |

where would the model suggest the new optimum to be located? What is the predicted value of the response at this location?

- (d) If the experimenter was interested in considering expanding the region to the range described in (c), how would you recommend proceeding with validating the results predicted by the model?

**8.43** Reconsider the experiment in Exercise 8.42 and Table E8.2. The experimenters selected a Box–Behnken design for their experiment.

- (a) This design has all of the design points that are not center runs at the same distance from the center of the design space in the scaled factors. How far are the design points from the center?
- (b) If the experimenters want to again use a spherical design with points the same distance from the center of carbon substrate concentration = 37.5 g/liter, broth pH = 3.5, fermentation temperature = 30°C, list the runs for a central composite design of the same design size.



TABLE E8.2 Data for Exercises 8.42 and 8.43

| X1 | X2 | X3 | Concentration |
|----|----|----|---------------|
| 1  | 1  | 0  | 20.47         |
| 1  | -1 | 0  | 24.48         |
| -1 | -1 | 0  | 15.69         |
| -1 | 0  | -1 | 15.87         |
| 0  | -1 | -1 | 23.45         |
| 0  | 0  | 0  | 15.26         |
| 0  | 0  | 0  | 14.99         |
| 0  | 1  | 1  | 36.42         |
| -1 | 0  | 1  | 13.68         |
| -1 | 1  | 0  | 15.37         |
| 1  | 0  | -1 | 30.72         |
| 1  | 0  | 1  | 14.73         |
| 0  | 1  | -1 | 12.05         |
| 0  | 0  | 0  | 16.27         |
| 0  | -1 | 1  | 12.2          |

(c) Compare the scaled prediction variance at the center and edges of the design region for the original design and the in (b). Which design do you prefer? Why?

**8.44** For a definitive screening design for six factors ( $X_1$ – $X_6$ ), with a total of 14 runs in the design space  $[-1,+1]^6$ ,

(a) List the design matrix.

(b) Create a fraction of design space plot for the design assuming a first order model for the six factors.

(c) Create a fraction of design space plot assuming the following model:

$$y = \beta_0 + \sum_{i=1}^6 \beta_i X_i + \beta_{12} X_1 X_2 + \beta_{14} X_1 X_4 + 6X_1 X_6 + \epsilon$$

(d) Create a fraction of design space plot assuming the following model:

$$y = \beta_0 + \sum_{i=1}^6 \beta_i X_i + \beta_{12} X_1 X_2 + \beta_{14} X_1 X_4 + \beta_{14} X_1 X_6 + \beta_{11} X_1^2 + \epsilon$$

(e) Compare the results from (b), (c), and (d) to summarize the effect of an increasingly complex model on scaled prediction variance.

(f) Compare the results from (d) to the fraction of design space plot for a six-factor face-centered CCD in the same range. Describe the trade-offs between the designs.

### Short Answer Questions

**8.45** Answer the following short answer questions:

(a) Why is it important to consider multiple aspects of the goodness of the design when selecting which experimental runs to choose?

- (b) Justify why it is a good strategy to not assuming a single chosen model is correct.
- (c) Describe the bias-variance trade-off and why it is important to not choose too small a model or too large a model.
- (d) Why is orthogonality an important property when choosing a designed experiment?
- (e) What are the required properties of a design to be variance-optimal for a first order model?
- (f) Describe how the structure of a central composite design allows estimation of the different terms in a second order model? Which design locations help to estimate which terms?
- (g) Why is rotatability a helpful property for a designed experiment?
- (h) Describe situations where it would be helpful to consider the scaled prediction variance? The unscaled prediction variance? The estimated prediction variance?
- (i) Describe how you would decide between a spherical and cuboidal design region for a particular experiment?
- (j) What are some of the trade-offs between a central composite design and a definitive screening design when the assumed model might be second-order?
- (k) Why is it important to incorporate orthogonality into the selection of the design's blocks?

**8.46** Answer the following True/False questions:

- (a) The 8-run  $2^3$  factorial design is a D-optimal design for fitting the first-order model with two-factor interaction terms included.
- (b) The face-centered cube is a rotatable design for fitting the second-order response surface model.
- (c) All second-order rotatable designs have the design points (apart from center points) the same distance from the design center.
- (d) In the face-centered cube, at least 3–5 center points are required to stabilize the prediction variance over a large portion of the design region.
- (e) A  $2^3$  factorial with four center points is variance-optimal for fitting the first-order model in three variables.
- (f) An experimenter must use at least one center run in a face-centered cube design so that he or she can solve the least squares normal equations to fit the second-order model.
- (g) The variance of prediction contours for a rotatable central composite design are circular even if the second-order model doesn't include all of the terms.
- (h) The necessary and sufficient conditions for rotatability for a  $d$ th order design depend on the design moments up through order  $2d$ .
- (i) An experimenter should use at least six center runs in a response surface design so that he or she will get a reliable estimate of experimental error.
- (j) A  $2^3$  factorial with two run at each corner of the cube is a variance-optimal design for fitting the first-order model in three variables.
- (k) Box–Benken designs are orthogonal for a first order model with two-way interactions.

- (l) The moment matrix for a design used to fit a model of order  $d$  contains moments up through order  $2d$ .
- (m) All central composite designs can be run in three blocks.
- (n) Rotatable and spherical central composite designs are always the same.

**8.47** Circle the correct answer for each question:

- (a) In a rotatable CCD the variance at the center of the design is
  - (i) Is always greater than the variance of the predicted response at the perimeter of the design.
  - (ii) Is always less than the variance of the predicted response at the perimeter of the design.
  - (iii) Is controlled entirely by the value of  $\alpha$
  - (iv) Depends mostly on the number of center runs
- (b) You have used a  $2^4$  factorial to fit a first-order model. The variance of an observation is  $\sigma^2 = 1$ . The variance of the predicted response at any corner of the design is equal to
  - (i) 4
  - (ii) 8
  - (iii) 12
  - (iv) None of the above
- (c) Consider a rotatable CCD in for 2 variables. The design is run two blocks. The factorial block has four center runs. To block orthogonally the number of center runs in the axial block should be
  - (i) 3
  - (ii) 2
  - (iii) 6
  - (iv) None of the above.
- (d) Consider a CCD in for four variables. The design is run in two blocks with 2 center runs in the factorial block and four center runs in the axial block. The value of  $\alpha$  that makes the design block orthogonally is
  - (i) 2
  - (ii)  $4/\sqrt{3}$
  - (iii)  $5/\sqrt{3}$
  - (iv) None of the above



---

# 9

---

## EXPERIMENTAL DESIGNS FOR FITTING RESPONSE SURFACES—II

In Chapter 8 we dealt with standard second-order response surface designs. The central composite designs (CCDs) and Box-Behnken designs (BBDs) are extremely popular with practitioners, and definitive screening designs (DSDs) provide a flexible alternative if fitting a second-order model may be desired when run as part of a screening design. The popularity of these standard designs is linked to the list of 11 important properties of response surface designs presented in Section 8.1. The CCD and BBD rate quite well for the 11 properties listed, particularly when they are augmented with center runs as recommended. One of the most important features associated with these two designs deals with run size. The run size is large enough to provide a comfortable margin for lack of fit, but not so large as to involve wasted degrees of freedom or unnecessary experimental expense.

Despite the importance of the CCD and the BBD, there are instances in which the researcher cannot afford the required number of runs. As a result, there is certainly a need for design classes that are either **saturated** or **near-saturated**, in other words, second-order designs that contain close to (but not less than)  $p$  design points where

$$p = 1 + 2k + \frac{k(k-1)}{2} \quad (9.1)$$

Note that  $p$  is the number of terms in the second-order model with 1 intercept,  $k$  first-order,  $k$  pure quadratic, and  $k(k-1)/2$  interaction terms. There are several design classes that allow saturated or near-saturated second-order designs. These are designs that should not be used unless cost prohibits the use of one of the standard designs. In a later section we make some comparisons that shed light on this point. In the section that follows we introduce a few of these design classes and make reference to others.



$$\mathbf{X}'\mathbf{X}_{\mathbf{D}_6} = \begin{matrix} & x_1 & x_2 & x_3 & x_1^2 & x_2^2 & x_3^2 & x_1x_2 & x_1x_3 & x_2x_3 \\ \left[ \begin{array}{cccccccccc} 13 & 0 & 0 & 0 & 10 & 10 & 10 & 0 & 0 & 0 \\ & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ & & 10 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ & & & 10 & 0 & 0 & 0 & -1 & 0 & 0 \\ & & & & 10 & 8 & 8 & 0 & 0 & -1 \\ & & & & & 10 & 8 & 0 & -1 & 0 \\ & & & & & & & 10 & -1 & 0 \\ & & & & & & & & 8 & -1 \\ & & & & & & & & & 8 \\ & & & & & & & & & & 8 \end{array} \right]
 \end{matrix}$$

**Hoke Designs for  $k > 3$**  The Hoke designs exist for any number of factors, with the  $\mathbf{D}_1$  to  $\mathbf{D}_3$  designs being saturated, and the  $\mathbf{D}_4$  to  $\mathbf{D}_7$  designs near-saturated. We can construct the designs using symmetry for a given patterns of levels. For example, for  $k = 4$  the saturated  $\mathbf{D}_2$  design is constructed by taking all combinations of the following combinations of the three levels:  $(-1, -1, -1, -1)$ ,  $(1, 1, 1, -1)$ ,  $(1, 1, -1, -1)$ , and  $(-1, 0, 0, 0)$ . This gives 1, 4, 6, and 4 rows for the design matrix, respectively. The  $\mathbf{D}_6$  design is the same as the  $\mathbf{D}_2$  design but with the four combinations of  $(1, 1, 1, 0)$  added.

$$\mathbf{D}_2 = \begin{matrix} & x_1 & x_2 & x_3 & x_4 \\ \left[ \begin{array}{cccc} -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{array} \right]
 \end{matrix}$$

The Hoke designs are an important class of designs, which offer a good option if a saturated or near-saturated design is required. In general, the use of saturated designs can be potentially risky as no degrees of freedom are available to estimate pure error or lack-of-fit. Hence if the assumed model is not sufficiently rich, detecting this is not possible. However, it is inevitable that these designs will be used where cost constraints are a major obstacle. In these cases, the Hoke designs are a good choice, and can typically be easily implemented since only three levels for each factor are required.

### 9.1.2 Koshal Design

Another class of design requiring a small run size is the family of **Koshal designs**. Koshal (1933) introduced this type of design for use in an effort to solve a set of likelihood equations. The designs are saturated for modeling of any response surface of order  $d$  ( $d = 1, 2, \dots$ ). We will not supply general details but will list Koshal designs for  $d = 1$  and 2.

**Koshal Design for First-Order Model** For the first-order model the Koshal design is simply the *one-factor-at-a-time* design. For  $k$  design variables we simply have

$$\mathbf{D} = \begin{array}{cccc} & x_1 & x_2 & \dots & x_k \\ \begin{bmatrix} \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \end{array}$$

Thus, for the special case in which three variables are of interest we have

$$\mathbf{D} = \begin{array}{ccc} & x_1 & x_2 & x_3 \\ \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

As a result, four coefficients can be estimated in the model

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i$$

It is important to note that the Koshal design for the first-order model does not allow estimation of any interaction terms.

**First-Order Plus Interaction** The Koshal family can be nicely extended to accommodate the first-order model with interaction. One simply augments the first-order Koshal design with **interaction rows**. For example, the appropriate design for  $k = 3$  is given by

$$\mathbf{D} = \begin{array}{ccc} & x_1 & x_2 & x_3 \\ \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \end{array}$$



**Second-Order Koshal Design** The Koshal design for fitting a second-order model must, of course, include at least three levels. The one-factor-at-a-time idea remains the basis for the design. For  $k=3$ , the design is given by

$$\mathbf{D} = \begin{array}{c} x_1 \quad x_2 \quad x_3 \\ \left[ \begin{array}{ccc} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{array} \right] \end{array}$$

The design contains three levels and 10 design points. Another form of the second-order Koshal design is given by

$$\mathbf{D} = \begin{array}{c} x_1 \quad x_2 \quad x_3 \\ \left[ \begin{array}{ccc} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{array} \right] \end{array}$$

It should be quite clear to the reader how to extend the Koshal designs to more than three variables.

### 9.1.3 Hybrid Designs

Roquemore (1976) developed a set of saturated or near-saturated second-order designs called **hybrid designs**. These hybrid designs are very efficient. They were created via an imaginative idea that involves the use of a central composite design for  $k-1$  variables, and the levels of the  $k$ th variable are determined in such a way as to create certain symmetries in the design. The result is a class of designs that are economical and either rotatable or

near-rotatable for  $k=3, 4, 6,$  and  $7$ . For example, for  $k=3$  and  $N=10$  we have the **hybrid 310** with design matrix

$$\mathbf{D}_{310} = \begin{matrix} & x_1 & x_2 & x_3 \\ \begin{bmatrix} 0 & 0 & 1.2906 \\ 0 & 0 & -0.1360 \\ -1 & -1 & 0.6386 \\ 1 & -1 & 0.6386 \\ -1 & 1 & 0.6386 \\ 1 & 1 & 0.6386 \\ 1.736 & 0 & -0.9273 \\ -1.736 & 0 & -0.9273 \\ 0 & 1.736 & -0.9273 \\ 0 & -1.736 & -0.9273 \end{bmatrix} \end{matrix}$$

It is important to note that the efficiency of the hybrid is based to a large extent on the fact that eight of the ten design points and two center runs on  $x_1$  and  $x_2$  are a central composite in these two variables. Then four levels are chosen on the remaining variable,  $x_3$ , with these levels chosen to *make odd moments zero, all second pure moments equal, and a near-rotatable design*. The design name, 310, comes from  $k=3$  and 10 distinct design points.

Roquemore developed two additional  $k=3$  hybrid designs. The design matrices are given by

$$\mathbf{D}_{311A} = \begin{matrix} & x_1 & x_2 & x_3 \\ \begin{bmatrix} 0 & 0 & \sqrt{2} \\ 0 & 0 & -\sqrt{2} \\ -1 & -1 & 1/\sqrt{2} \\ 1 & -1 & 1/\sqrt{2} \\ -1 & 1 & 1/\sqrt{2} \\ 1 & 1 & 1/\sqrt{2} \\ \sqrt{2} & 0 & -1/\sqrt{2} \\ -\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & \sqrt{2} & -1/\sqrt{2} \\ 0 & \sqrt{2} & -1/\sqrt{2} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \end{matrix}$$

$$\mathbf{D}_{311B} = \begin{matrix} & x_1 & x_2 & x_3 \\ \begin{bmatrix} 0 & 0 & \sqrt{6} \\ 0 & 0 & -\sqrt{6} \\ -0.7507 & 2.1063 & 1 \\ 2.1063 & 0.7507 & 1 \\ 0.7507 & -2.1063 & 1 \\ -2.1063 & -0.7507 & 1 \\ 0.7507 & 2.1063 & -1 \\ 2.1063 & -0.7507 & -1 \\ -0.7507 & -2.1063 & -1 \\ -2.1063 & 0.7507 & -1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \end{matrix}$$

Both the 311A and 311B are near-rotatable. Eleven design points include a center run to avoid near-singularity.

There are three hybrid designs for  $k = 4$ . They are given by

$$\mathbf{D}_{416A} = \begin{matrix} & x_1 & x_2 & x_3 & x_4 \\ \begin{bmatrix} 0 & 0 & 0 & 1.7844 \\ 0 & 0 & 0 & -1.4945 \\ -1 & -1 & -1 & 0.6444 \\ 1 & -1 & -1 & 0.6444 \\ -1 & 1 & -1 & 0.6444 \\ 1 & 1 & -1 & 0.6444 \\ -1 & -1 & 1 & 0.6444 \\ 1 & -1 & 1 & 0.6444 \\ -1 & 1 & 1 & 0.6444 \\ 1 & 1 & 1 & 0.6444 \\ 1.6853 & 0 & 0 & -0.9075 \\ -1.6853 & 0 & 0 & -0.9075 \\ 0 & 1.6853 & 0 & -0.9075 \\ 0 & -1.6553 & 0 & -0.9075 \\ 0 & 0 & 1.6853 & -0.9075 \\ 0 & 0 & -1.6853 & -0.9075 \end{bmatrix} \end{matrix}$$

$$\mathbf{D}_{416B} = \begin{matrix} & x_1 & x_2 & x_3 & x_4 \\ \begin{bmatrix} 0 & 0 & 0 & 1.7317 \\ 0 & 0 & 0 & -0.2692 \\ \pm 1 & \pm 1 & \pm 1 & \mathbf{0.6045} \\ \pm 1.5177 & \mathbf{0} & \mathbf{0} & -\mathbf{1.0498} \\ \mathbf{0} & \pm 1.5177 & \mathbf{0} & -\mathbf{1.0498} \\ \mathbf{0} & \mathbf{0} & \pm 1.5177 & -\mathbf{1.0498} \end{bmatrix} \end{matrix}$$

$$\mathbf{D}_{416C} = \begin{matrix} & x_1 & x_2 & x_3 & x_4 \\ \begin{bmatrix} 0 & 0 & 0 & 1.7658 \\ 0 & 0 & 0 & 0 \\ \pm 1 & \pm 1 & \pm 1 & \mathbf{0.5675} \\ \pm 1.4697 & \mathbf{0} & \mathbf{0} & -\mathbf{1.0509} \\ \mathbf{0} & \pm 1.4697 & \mathbf{0} & -\mathbf{1.0509} \\ \mathbf{0} & \mathbf{0} & \pm 1.4697 & -\mathbf{1.0509} \end{bmatrix} \end{matrix}$$

We have condensed the notation for 416B and 416C. The notation  $\pm 1$  indicates a  $2^3$  factorial in  $x_1, x_2,$  and  $x_3$  with  $x_4$  fixed. The design points following the  $2^3$  factorial are axial points in  $x_1, x_2,$  and  $x_3$  with  $x_4$  fixed. A center run is used in 416C to avoid near-singularity. The notation  $\pm 1.4697$  or  $\pm 1.5177$  implies two axial points. When possible, one or two more center runs should be added to the hybrid design.

There are two hybrid designs for six design variables. Both include 28 runs, and thus both are saturated. The 628A design is as follows:

$$D_{628A} = \begin{matrix} & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ \begin{matrix} 0 \\ \pm 1 \\ \pm 2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ \pm 1 \\ 0 \\ \pm 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ \pm 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ \pm 1 \\ 0 \\ 0 \\ \pm 2 \\ 0 \\ 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ \pm 1 \\ 0 \\ 0 \\ 0 \\ \pm 2 \\ 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ \pm 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ \pm 2 \\ 0 \end{matrix} & \begin{matrix} 4/\sqrt{3} \\ 1/\sqrt{3} \\ -2/\sqrt{3} \\ -2/\sqrt{3} \\ -2/\sqrt{3} \\ -2/\sqrt{3} \\ -2/\sqrt{3} \\ 0 \end{matrix} \end{matrix}$$

The notation  $\pm 1$  indicates a resolution  $V 2^{5-1}$  in  $x_1, x_2, x_3, x_4,$  and  $x_5$ , while  $x_6$  is held  $1/\sqrt{3}$ . The  $\pm 2$  implies two axial points. A second six-variable design, the 628B, is constructed in a manner similar to that of 628A. The components of the design are as follows:

1. Resolution  $V 2^{5-1}$  in  $x_1 - x_5$  with  $x_6 = 0.6096$
2. Axial points in  $x_1 - x_5$  with  $\alpha = 2.1749$  and  $x_6 = -1.0310$
3. Two additional center points with  $x_6 = 2.397$  and  $-1.8110$

**General Comments Concerning the Hybrid Designs** As we indicated earlier, the hybrid designs represent a very efficient class of saturated or near-saturated second-order designs. The hybrid designs are quite competitive with central composite designs *when the criteria for comparison takes design size into account*—for example, when one makes comparisons using  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ . Further discussion of design comparisons is in Sections 9.2 and 9.3.

It has been our experience that hybrid designs are not used as much in industrial applications as they should be. As previously discussed, the use of saturated or near-saturated response surface designs should be avoided. However, where cost constraints are major obstacles, the hybrid design is a good choice. It is likely that potential users are reluctant to use the hybrid designs because of the “messy levels” required of the extra design variable—that is, the variable not involved in the central composite structure. One must remember that the levels reported in the design matrices in this text are those solved for by Roquemore in order to achieve certain ideal design conditions. An efficient design will still result if one merely approximates these levels. *Minor errors in control* do not alter the efficiencies of these designs. One should assign the extra design variable to a factor that is easiest to alter and thus accommodates these levels. Of course, many practitioners are attracted to designs that involve three evenly spaced levels or levels that provide little difficulty when the experiment is conducted.

#### 9.1.4 The Small Composite Design

The small composite design (Draper, 1985) is commonly available in some software packages, and is present here because of its popularity. However, we would like to discourage the use of this design because of its poor estimation and prediction performance. More details about its performance are given in Section 9.2. The small composite design gets its

name from the ideas of the central composite, but the factorial portion is neither a complete  $2^k$  nor a resolution V fraction, but, rather, a special resolution III fraction in which no four-letter word is among the defining relations. This type of fraction is often called *resolution III\**. As a result, the total run size is reduced from that of the CCD—hence the term **small composite design**. For  $k = 3$ , we have

$$\mathbf{D} = \begin{matrix} & x_1 & x_2 & x_3 \\ \begin{bmatrix} -1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \\ -\alpha & 0 & 0 \\ \alpha & 0 & 0 \\ 0 & -\alpha & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & -\alpha \\ 0 & 0 & \alpha \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} & & \end{matrix} \tag{9.2}$$

The factorial portion is the fraction generated using  $I = -ABC$ . Obviously, the alternative fraction would be equally satisfactory in the construction of the small composite design. In this case,  $p$ , the number of second-order parameters, is 10, and thus the design is one degree of freedom above saturation. Multiple center runs allow degrees of freedom for pure error but if the fitted model is second order, there will be one degree of freedom for lack of fit.

If we focus on the example design in Equation 9.2, it becomes obvious that in the factorial portion linear main effect terms are aliased with two-factor interaction terms. Hartley (1959) observed that in spite of this, all coefficients in the second-order model are estimable because the linear coefficients benefit from the axial points, though axial points provide no information on the estimation of interaction coefficients. One may gain some additional insight into this by observing the matrix  $\mathbf{X}'\mathbf{X}$  for the  $k = 3$  small composite design (SCD) contrasted with that for the  $k = 3$  full central composite (assume a single center run for both designs):

$$\mathbf{X}'\mathbf{X}_{\text{SCD}} = \begin{matrix} & x_1 & x_2 & x_3 & x_1^2 & x_2^2 & x_3^2 & x_1x_2 & x_1x_3 & x_2x_3 \\ \begin{bmatrix} 11 & 0 & 0 & 0 & 4 + 2\alpha^2 & 4 + 2\alpha^2 & 4 + 2\alpha^2 & 0 & 0 & 0 \\ & 4 + 2\alpha^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -4 \\ & & 4 + 2\alpha^2 & 0 & 0 & 0 & 0 & 0 & -4 & 0 \\ & & & 4 + 2\alpha^2 & 0 & 0 & 0 & -4 & 0 & 0 \\ & & & & 4 + 2\alpha^4 & 4 & 4 & 0 & 0 & 0 \\ & & & & & 4 + 2\alpha^4 & 4 & 0 & 0 & 0 \\ & & & & & & 4 + 2\alpha^4 & 0 & 0 & 0 \\ & & & & & & & 4 & 0 & 0 \\ & & & & & & & & 4 & 0 \\ & & & & & & & & & 4 \end{bmatrix} & & \end{matrix} \tag{9.3}$$

$$\mathbf{X}'\mathbf{X}_{\text{CCD}} = \begin{bmatrix} 15 & 0 & 0 & 0 & 8 + 2\alpha^2 & 8 + 2\alpha^2 & 8 + 2\alpha^2 & 0 & 0 & 0 \\ 0 & 8 + 2\alpha^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 8 + 2\alpha^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 + 2\alpha^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 8 + 2\alpha^2 & 0 & 0 & 0 & 8 + 2\alpha^4 & 8 & 8 & 0 & 0 & 0 \\ 8 + 2\alpha^2 & 0 & 0 & 0 & 8 & 8 + 2\alpha^4 & 8 & 0 & 0 & 0 \\ 8 + 2\alpha^2 & 0 & 0 & 0 & 8 & 8 & 8 + 2\alpha^4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 + 2\alpha^4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 \end{bmatrix} \tag{9.4}$$

The distinction between the two designs stems from the fact that in the CCD all linear main effects and two-factor interactions are mutually orthogonal, whereas in the SCD the  $-4$  values reflect the nonorthogonality of  $x_1$  with  $x_2x_3$ ,  $x_2$  with  $x_1x_3$ , and  $x_3$  with  $x_1x_2$ . The correlation among these model terms has a strong effect on the variance of the regression coefficients of  $x_1, x_2, x_3, x_1x_2, x_1x_3$ , and  $x_2x_3$ . In fact, a comparison can be made by creating scaled variances for the two designs, that is,  $N \text{Var}(b_i)/\sigma^2$  for  $i = 1, 2, 3$  and  $N \text{Var}(b_{ij})/\sigma^2$  for  $i, j = 1, 2, 3; i \neq j$ . Table 9.1 shows these values taken from appropriate diagonals of the matrix  $N(\mathbf{X}'\mathbf{X})^{-1}$ . In our illustration,  $\alpha$  was taken to be  $\sqrt{3}$  and  $n_c = 3$  for each design. Notice that the designs are quite similar for estimation of pure quadratic coefficients. However, the small composite design suffers considerably in efficiency for estimation of linear and interaction coefficients.

The efficiencies of model coefficients certainly suggest that the small composite design should not be used, as several of the other designs, such as the Hoke and hybrid designs have similar size and perform much better. Coefficient by coefficient comparisons of variances is certainly not the only, nor necessarily even the best, method for comparing designs. Graphical methods for comparing designs in terms of scaled prediction variance (i.e.,  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ ) can be used. These methods are presented later in this chapter, and the comparison between the CCD and SCD will be displayed prominently.

**Small Composite Design for  $k = 2$**  For  $k = 2$  there is a small composite design in which the factorial portion is the one-half fraction of a  $2^2$ . The design is given by

**TABLE 9.1 Scaled Variances of Model Coefficients for CCD and SCD**

|                   | $b_0$  | $b_i$  | $b_{ii}$ | $b_{ij}$ |
|-------------------|--------|--------|----------|----------|
| CCD ( $N = 17$ )  | 5.6666 | 1.2143 | 1.3942   | 2.125    |
| SCD ( $AC = 13$ ) | 4.3333 | 2.1666 | 1.1074   | 5.4166   |

$$\mathbf{D} = \begin{matrix} & x_1 & x_2 \\ \begin{bmatrix} -1 & -1 \\ 1 & 1 \\ -\alpha & 0 \\ \alpha & 0 \\ 0 & -\alpha \\ 0 & \alpha \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \end{matrix}$$

The defining relation for the factorial portion is given by  $I=AB$ . With six parameters to estimate, this design affords one degree of freedom for lack of fit. For  $k=2$  the hexagon is a more efficient design. This SCD design may find some use if the design region is cuboidal, in which case  $\alpha$  must be 1.0.

**Small Composite Designs for  $k > 3$**  The general SCD is a resolution III\* fraction of a  $2^k$  augmented with axial points and runs in the center of the design. This structure applies for all  $k > 2$ . For example, for  $k=4$  a SCD is given by

$$\mathbf{D} = \begin{matrix} & x_1 & x_2 & x_3 & x_4 \\ \begin{bmatrix} -1 & -1 & -1 & +1 \\ 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ -\alpha & 0 & 0 & 0 \\ \alpha & 0 & 0 & 0 \\ 0 & -\alpha & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & -\alpha & 0 \\ 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & -\alpha \\ 0 & 0 & 0 & -\alpha \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \end{matrix} \tag{9.5}$$

The factorial portion of the design in Equation 9.5 is a one-half fraction of a  $2^4$  with  $I=ABD$  as the defining relation. The design points yield 17 degrees of freedom for parameter estimation. Thus, the design is not saturated but results in only two degrees of freedom for lack of fit. In comparison, the CCD for  $k=4$  allows 10 degrees of freedom for lack of fit. Comparisons between scaled variances of coefficients of the linear and interaction coefficients of the SCD with that of the CCD essentially tell the same story as in the  $k=3$  case. The superiority of the CCD is quite apparent, and indeed other small designs also outperform the SCD substantially.

An additional important point should be made regarding the  $k=4$  case. One is tempted to use a resolution IV fraction rather than resolution III\* for  $k=4$  and possibly for larger values of  $k$  where higher fractions might be exploited. For example, for  $k=6$  a resolution

IV  $2^{6-2}$  fraction is available. This allows 16 factorial points plus 12 axial points plus center runs, and the result is one degree of freedom over saturation. However, the design *cannot* involve a resolution IV fraction: the result would be complete aliasing among two-factor interactions. One must be mindful of the fact that unlike linear or pure quadratic coefficients, all information on interaction coefficients is derived from the factorial portion. Thus, aliasing among two-factor interaction terms in the factorial portion results in aliasing of two-factor interactions for the entire design.

***Further Comments on the Small Composite Design*** Despite its availability in software packages, the small composite design is typically not a good choice when cost constraints restrict the design size to substantially smaller than that required by the central composite or Box–Behnken designs. If the primary goal of the experiment is to understand the second-order model, other alternatives such as the Hoke or hybrid designs provide better estimation and prediction with a similar small sample size.

One potential virtue of the SCD may arise in sequential experimentation. Most of the discussion regarding sequential experimentation has centered on the transition from the first-order fitted model to the second-order model. One should regard response surface methodology (RSM) as nearly always involving iterative experimentation and iterative decision-making. A first-stage experiment with an SCD may suggest that the current region is, indeed, the proper one. That is, there are no ridge conditions and no need to move to an alternative region. Then a second-stage augmentation of the factorial portion is conducted; the result is now either a full factorial or resolution V fraction that accompanies the axial points. As a result, the second stage is more efficient than what was used in the first stage. Further information about the SCD and comparisons with other designs will be given later in the chapter.

### 9.1.5 Some Saturated or Near-Saturated Cuboidal Designs

During the 1970s and early 1980s, considerable attention was given to the development of efficient second-order designs on a cube. This movement settled down considerably in the 1980s, and attention shifted to developments that led to general computer-generated experimental designs. Computer-generated design will be discussed in general in Section 9.3. However, there remain several interesting design classes for a cube that have potential use. Designs developed by Notz (1982), and Box and Draper (1974) are particularly noteworthy. As an example, the Notz three-variable design involves seven design points from a  $2^3$  factorial plus three one-factor-at-a-time axial points. Specifically, this design is given by

$$\mathbf{D} = \begin{array}{c} x_1 \quad x_2 \quad x_3 \\ \left[ \begin{array}{ccc} -1 & -1 & -1 \\ 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right] \end{array}$$



A design consisting of one additional point, namely (+1, +1, +1) to complete the 2<sup>3</sup> factorial, resulting in 11 design points, is a nice augmentation of the three-factor Notz design. The four-variable Notz design is constructed in a similar manner, with 11 points from the 2<sup>4</sup> factorial plus four axial points.

**9.1.6 Equiradial Designs**

There are some special and interesting two-factor designs that serve as alternatives to the central composite design. They are the class of **equiradial designs**, beginning with a pentagon (five equally spaced points on a circle); the pentagon, hexagon, heptagon, and so on, do require center runs because they are designs on a common sphere and, as we shall show, are rotatable.

The design matrix for the equiradial design can be written

$$[\rho \cos(\theta + 2\pi u/n_1), \quad \rho \sin(\theta + 2\pi u/n_1)], \quad u = 0, 1, 2, \dots, n_1 - 1 \quad (9.6)$$

where  $\rho$  is the design radius. Here,  $n_1$  is the number of points on the sphere, say, 5, 6, 7, 8, ... . In addition to the  $n_1$  points indicated by Equation 9.6, we will assume  $n_c$  center runs. The value that one chooses for  $\rho$  merely determines the nature of the scaling. It turns out that the choice of  $\theta$  has no effect on  $\mathbf{X}'\mathbf{X}$ ; that is, it has no effect on the design as far as the variance structure is concerned. As a result, all of the equiradial second-order designs (i.e., using  $n_1 = 5, 6, 7, 8, \dots$ ) are such that the matrix  $\mathbf{X}'\mathbf{X}$  is *invariant to design rotation*.

A very useful and interesting special case is the **hexagon**—that is, six equally spaced points on a circle. Thus  $n_1 = 6$ ; setting  $\theta = 0$  and  $\rho = 1$ , we have the following design matrix using  $n_c = 3$ :

$$\mathbf{D} = \begin{matrix} & \begin{matrix} x_1 & x_2 \end{matrix} \\ \begin{bmatrix} 1 & 0 \\ 0.5 & \sqrt{0.75} \\ -0.5 & \sqrt{0.75} \\ -1 & 0 \\ -0.5 & -\sqrt{0.75} \\ 0.5 & -\sqrt{0.75} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} & \end{matrix} \quad (9.7)$$

Note that  $x_1$  is at five levels and  $x_2$  is at three levels for this particular rotation. Figure 9.1 shows the design. Note the six evenly spaced points.

It is of interest to consider the important design moments for the hexagon—that is, the moments through order four. From Equation 9.7 we have, for the hexagon,

$$[i] = [ij] = [iij] = [iii] = [iiij] = 0, \quad i \neq j = 1, 2 \text{ (odd moments zero)}$$

$$[iiii] = \frac{9}{4N}, \quad [iijj] = \frac{3}{4N}, \quad i \neq j = 1, 2$$

Because all design moments through order four are zero and  $[iiii]/[iijj] = 3$ , the hexagon is indeed rotatable. The reader should also note that these conditions are independent of  $\rho$  and  $\theta$ .

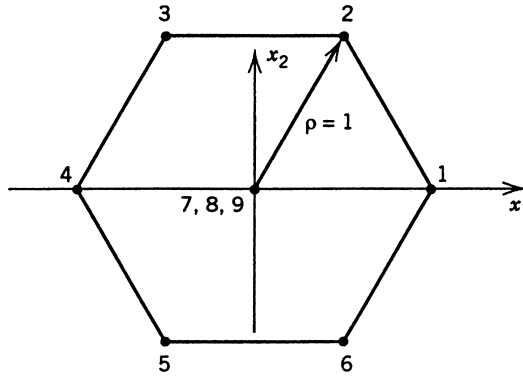


Figure 9.1 Design points for the hexagon.

It turns out that all equiradial designs for  $n_1 \geq 5$  are rotatable. Appendix 2 shows a development of the moment conditions. The relevant moments for the equiradial designs are as follows:

1. All odd moments through order four are zero.
2. The key even moments are

$$\begin{aligned}
 [ii] &= \frac{\rho^2 n_1}{2N}, \quad i = 1, 2 \\
 [iiii] &= \frac{3\rho^4 n_1}{8N}, \quad i = 1, 2 \\
 [ijij] &= \frac{\rho^4 n_1}{8N}, \quad i \neq j = 1, 2
 \end{aligned}$$

Of course, these conditions imply that all equiradial designs for  $n_1 \geq 5$  are rotatable.

**Special Case of The Equiradial Design—The CCD for  $k=2$**  For  $n_1 = 8$ , the equiradial design is the **octagon**. As one might expect, the use of Equation 9.6 for  $n_1 = 8$  produces (using  $\rho = \sqrt{2}$ )

$$\begin{matrix}
 x_1 & x_2 \\
 \left[ \begin{array}{cc}
 -1 & -1 \\
 -1 & 1 \\
 1 & -1 \\
 1 & 1 \\
 \sqrt{2} & 0 \\
 -\sqrt{2} & 0 \\
 0 & -\sqrt{2} \\
 0 & \sqrt{2} \\
 \mathbf{0} & \mathbf{0}
 \end{array} \right]
 \end{matrix}$$

which, of course, is the  $k=2$  rotatable CCD.

**Summary Comments on the Equiradial Designs** The equiradial designs are an interesting family of designs that enjoy some usage in the case where only two design variables are involved. One should use the CCD (the octagon) whenever possible. However, there are situations when cost constraints do not allow the use of the octagon. The pentagon (plus center runs) is a saturated design and thus should not be used unless absolutely necessary. The hexagon is a nice design that allows one degree of freedom for lack of fit. The heptagon also has its place in some applications. The octagon, hexagon, and nonagon (nine equally spaced points) possess the added property of *orthogonal blocking*, which will be discussed in the next section.

We have already indicated that the equiradial designs require the use of center runs. Reasonable stability of  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  is achieved with two to four center runs.

## 9.2 GENERAL CRITERIA FOR CONSTRUCTING, EVALUATING, AND COMPARING DESIGNED EXPERIMENTS

As we indicated in Section 9.1, the 1980s began an era of using the computer for design construction. This period encouraged many new users to consider and use designed experiments, simply because it became easier to find appropriate designs. While computer-generated designs are flexible and easy to use, there are also negative aspects. One of the most important disadvantages is that users often treat the designs generated by a computer package like a black box, where an “optimal” design must be evaluated relative to the criterion that was used to optimize it and how closely the design matches the model identified by the user as appropriate to characterize the relationship between the input factors and the response. Another substantial disadvantage of computer-generated designs is their tendency to focus on just a single criterion. Recall from Chapter 8 that there are numerous characteristics typically associated with good designs, and focusing solely on a single objective can lead to poor performance on other important aspects of the design.

Much of what is available in evaluation and comparison of RSM designs as well as computer-generated design are the results of the work of Kiefer (1959, 1961) and Kiefer and Wolfowitz (1959) in **optimal design theory**. Their work is couched in a measure theoretic approach in which a designed experiment is viewed in terms of design measure. Design optimality moved into the practical arena in the 1970s and 1980s as designs were put forth as being efficient in terms of criteria inspired by Kiefer and his coworkers. Computer algorithms were developed that allowed “best” designs to be generated by a computer package based on the practitioner’s choice of sample size, model, ranges on variables, and other constraints. The notion of reducing the “goodness” of a design to a single number is often too ambitious, and so optimality should most often be considered as one of many aspects to be balanced. Recall from Section 8.1, that there are often inherent trade-offs between design criteria, which should be explicitly considered when selecting a design. For example, an optimal design for a particular model likely will not have good ability to evaluate the assumptions of the model or assess lack of fit. Hence, optimality should most often be considered as one of many aspects to be balanced. Later in this chapter, some of the ideas from multiple objective optimization in Chapter 7 will be introduced as a method to formally consider several design criteria.

There are three situations where using a computer-generated design based on a single criterion may be highly beneficial:

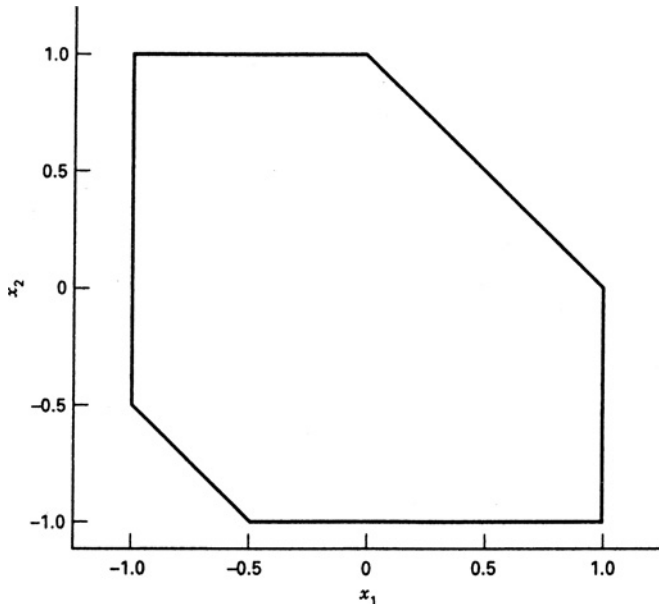
1. *An irregular experimental region.* If the region of interest for the experiment is not a cube or a sphere, standard designs may not be the best choice. Irregular regions of interest occur fairly often. For example, an experimenter is investigating the properties of a particular adhesive. The adhesive is applied to two parts and then cured at an elevated temperature. The two factors of interest are the amount of adhesive applied and the cure temperature. Over the ranges of these two factors, taken as  $-1$  to  $+1$  on the usual coded variable scale, the experimenter knows that if too little adhesive is applied and the cure temperature is too low, the parts will not bond satisfactorily. In terms of the coded variables, this leads to a **constraint** on the design variables, say

$$-1.5 \leq x_1 + x_2$$

where  $x_1$  represents the application amount of adhesive and  $x_2$  represents the temperature. Furthermore, if the temperature is too high and too much adhesive is applied, either the parts will be damaged by heat stress or an inadequate bond will result. Thus, there is another constraint on the factor levels:

$$x_1 + x_2 \leq 1$$

Figure 9.2 shows the experimental region that results from applying these constraints. Notice that the constraints effectively remove two corners of the square, producing an irregular experimental region. There is no standard response surface design that will fit exactly into this region.



**Figure 9.2** A constrained design region in two variables.

2. *A nonstandard model.* Usually an experimenter selects a first- or second-order response surface model, realizing that this **empirical model** is an approximation to the true underlying mechanism. However, sometimes the experimenter may have some special knowledge or insight about the process being studied that may suggest a nonstandard model. For example, the model

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{12}x_1x_2 + \beta_{11}x_1^2 + \beta_{22}x_2^2 \\ + \beta_{112}x_1^2x_2 + \beta_{1112}x_1^3x_2 + \varepsilon$$

may be of interest. The experimenter would be interested in obtaining an efficient design for fitting this reduced quartic model. As another illustration, sometimes we encounter response surface problems where some of the design factors are categorical variables. There may be no standard response surface designs for this situation. We will discuss this further in Section 9.3.2.

3. *Unusual sample size or block size requirements.* Occasionally an experimenter may need to reduce the number of runs required by a standard response surface design. For example, suppose we intend to fit a second-order model in four variables. The central composite design for this situation requires between 28 and 30 runs, depending on the number of center points selected. However, the model has only 15 terms. If the runs are extremely expensive or time-consuming, the experimenter may want a design with fewer trials. Although computer-generated designs can be used for this purpose, there may be better approaches. For example, the Hoke D6 design uses only 19 runs, has very good properties, and a hybrid design with as few as 16 runs is also available. These can be superior choices to using a computer-generated design to reduce the number of trials. Unusual requirements concerning either the number of blocks or the block size can also suggest the use of a computer-generated design.

### 9.2.1 Practical Design Optimality

Design optimality criteria are characterized by letters of the alphabet and, as a result, are often called **alphabetic optimality criteria**. Some criteria focus on good estimation of model parameters, while others focus on good prediction in the design region. The best-known and most often used criterion is *D*-optimality.

***D*-Optimality and *D*-Efficiency** *D*-optimality, which focuses on good model parameter estimation, is based on the notion that the designed experiment should be chosen so as to achieve certain properties in the **moment matrix**

$$\mathbf{M} = \frac{\mathbf{X}'\mathbf{X}}{N} \quad (9.8)$$

The reader recalls the importance of the elements of the moment matrix in the determination of rotatability. Also, the inverse of  $\mathbf{M}$ , namely

$$\mathbf{M}^{-1} = N(\mathbf{X}'\mathbf{X})^{-1} \quad (9.9)$$

(the **scaled dispersion matrix**), contains variances and covariances of the regression coefficients, scaled by  $N/\sigma^2$ . As a result, control of the moment matrix by design implies control of the variances and covariances.

It turns out that an important norm on the moment matrix is the **determinant**, that is,

$$|\mathbf{M}| = \frac{|\mathbf{X}'\mathbf{X}|}{N^p} \quad (9.10)$$

where  $p$  is the number of parameters in the model. Under the assumption of independent normal model errors with constant variance, the *determinant of  $\mathbf{X}'\mathbf{X}$  is inversely proportional to the square of the volume of the confidence region on the regression coefficients*. The volume of the confidence region is relevant because it reflects how well the set of coefficients are estimated. A small  $|\mathbf{X}'\mathbf{X}|$  and hence large  $|\mathbf{X}'\mathbf{X}|^{-1} = 1/|\mathbf{X}'\mathbf{X}|$  implies poor estimation of  $\boldsymbol{\beta}$  in the model. For the response surface model which is a special case of the general linear model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

the  $100(1 - \alpha)\%$  confidence ellipsoid on  $\boldsymbol{\beta}$  under the assumption  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$  is given by solutions to  $\boldsymbol{\beta}$  in

$$\frac{(\mathbf{b} - \boldsymbol{\beta})'(\mathbf{X}'\mathbf{X})(\mathbf{b} - \boldsymbol{\beta})}{ps^2} \leq F_{\alpha, p, n-p}$$

or

$$(\mathbf{b} - \boldsymbol{\beta})'(\mathbf{X}'\mathbf{X})(\mathbf{b} - \boldsymbol{\beta}) \leq C$$

where  $C = ps^2 F_{\alpha, p, n-p}$ . As a result, the volume of the confidence region is the volume of the ellipsoid

$$(\mathbf{b} - \boldsymbol{\beta})'(\mathbf{X}'\mathbf{X})(\mathbf{b} - \boldsymbol{\beta}) = C$$

Suppose we consider the orthogonal matrix  $\mathbf{Q}$  such that

$$\mathbf{Q}'(\mathbf{X}'\mathbf{X})\mathbf{Q} = \begin{bmatrix} \lambda_1^* & & & 0 \\ & \lambda_2^* & & \\ & & \ddots & \\ 0 & & & \lambda_p^* \end{bmatrix}$$

where the  $\lambda_i^*$  are the eigenvalues of  $\mathbf{X}'\mathbf{X}$ . We can write

$$\mathbf{X}'\mathbf{X} = \mathbf{Q} \begin{bmatrix} \lambda_i^* & & 0 \\ & \ddots & \\ 0 & & \lambda_p^* \end{bmatrix} \mathbf{Q}'$$

As a result, we can write

$$(\mathbf{b} - \boldsymbol{\beta})'(\mathbf{X}'\mathbf{X})(\mathbf{b} - \boldsymbol{\beta}) = \sum_{i=1}^p (b_i - \beta_i)^2 \lambda_i^* = C$$

The volume of this ellipsoid is proportional to  $[\prod_{i=1}^p (1/\lambda_i^*)]^{1/2}$ . The quantity  $[\prod_{i=1}^p (1/\lambda_i^*)]^{1/2} = 1/|\mathbf{X}'\mathbf{X}|^{1/2}$ . As a result, the volume of the ellipsoid is inversely proportional to the square root of the determinant of  $\mathbf{X}'\mathbf{X}$ .

A *D-optimal design* is one in which  $|\mathbf{M}| = |\mathbf{X}'\mathbf{X}|/N^p$  is maximized<sup>†</sup>; that is,

$$\text{Max}_{\zeta} |\mathbf{M}(\zeta)| \tag{9.11}$$

where Max implies that the maximum is taken over all designs  $\zeta$ . As a result, it is natural to define the *D-efficiency* of a design  $\zeta^*$  as

$$D_{\text{eff}} = \left( |\mathbf{M}(\zeta^*)| / \text{Max}_{\zeta} |\mathbf{M}(\zeta)| \right)^{1/p} \tag{9.12}$$

Here, the  $1/p$  power takes account of the  $p$  parameter estimates being assessed when one computes the determinant of the variance–covariance matrix. The definition of *D-efficiency* in Equation 9.12 allows for comparing designs that have different sample sizes by comparing *D-efficiencies*. Since it is helpful to balance the multiple objectives of a good design outlined in Section 8.1, *D-efficiency* can be a useful tool for quantifying the quality of our estimation of model parameters.

There have been many studies that have been conducted to compare standard experimental designs through the use of *D-efficiency*. Readers are referred to Lucas (1976) for comparisons among some of the designs that we have discussed in this chapter and in Chapter 8. St. John and Draper (1975) give an excellent review of *D-optimality* and provide much insight into pragmatic use of design optimality. One should also read Box and Draper (1971), Myers et al. (1989), and the text by Pukelsheim (1995).

Recall that in Chapter 8 we discussed the notion of variance-optimal designs in the first-order and first-order-plus-interaction case. While our main objective in this chapter is to continue to focus on second-order designs, it is natural to consider *D-optimality* and *D-efficiency* in the case of simpler models. Consider the variance-optimal design, namely the orthogonal design with all levels at the  $\pm 1$  extremes of the experimental region. The moment matrix is given by

$$\mathbf{M} = \frac{\mathbf{X}'\mathbf{X}}{N} = \mathbf{I}_p$$

It is not difficult to show that for the first-order and first-order-plus-interaction cases, optimality extends to the determinant, that is,

$$\text{Max}_{\zeta} |\mathbf{M}(\zeta)| = 1,$$

<sup>†</sup>Equivalently, one may minimize  $N^p |\mathbf{X}'\mathbf{X}|^{-1}$ .

and thus for these simpler models, the so-called variance-optimal design is also optimal in the determinant sense—that is,  $D$ -optimal.

Before we embark on some comparisons among standard designs and discussion of how  $D$ -optimality affects practical tools in the use of computer-generated designs, we offer a few other important optimality criteria and methods of comparison.

**$A$ -Optimality** The concept of  **$A$ -optimality**, which also aims to estimate model parameters well, deals with the individual variances of the regression coefficients. Unlike  $D$ -optimality, it does not consider the covariances among coefficients; recall that the variances of regression coefficients appear on the diagonals of  $(\mathbf{X}'\mathbf{X})^{-1}$ .  $A$ -optimality is defined as

$$\text{Min}_{\zeta} \text{tr}[\mathbf{M}(\zeta)]^{-1} \quad (9.13)$$

where  $\text{tr}$  represents trace, that is, the sum of the variances of the coefficients (weighted by  $N$ ). Some computer-generated design packages make use of  $A$ -optimality.

**Criteria Associated with Prediction Variance ( $G$ -,  $V$ -, and  $I$ -Optimality)** Throughout Chapter 8, we consistently made reference to the use of the scaled prediction variance  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2 = v(\mathbf{x})$  as an important measure of performance. The primary goal of many designed experiments is to allow for good prediction throughout the design space. Hence, focusing on  $v(\mathbf{x})$  makes direct assessment possible. Our feeling is that practical designers of experiments don't make use of this measure as they should. One clear disadvantage, of course, is that, unlike  $\text{tr}(\mathbf{M}^{-1})$  or  $|\mathbf{M}|$ ,  $v(\mathbf{x})$  is not a single number but rather depends on the location  $\mathbf{x}$  at which one is predicting. In fact, recall that

$$v(\mathbf{x}) = \frac{N \text{Var}[\hat{y}(\mathbf{x})]}{\sigma^2} = N \mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)}$$

where  $\mathbf{x}^{(m)}$  reflects location in the design space as well as the nature of the model. In previous discussions in Chapter 8, we focused on attempts to stabilize  $v(\mathbf{x})$  by proper design (choice of center runs).

One interesting design optimality criterion that focuses on  $v(\mathbf{x})$  is  $G$ -optimality.  **$G$ -optimality** and the corresponding  **$G$ -efficiency** emphasize the use of designs for which the *maximum  $v(\mathbf{x})$  in the region of the design is not too large*. It seeks to protect against the worst case prediction variance, since when we use the results of our analysis we may wish to predict new response values anywhere in the design region. A  $G$ -optimal design  $\zeta$  is one in which we have

$$\text{Min}_{\zeta} \left[ \text{Max}_{\mathbf{x} \in R} v(\mathbf{x}) \right]$$

Note that this is equivalent to

$$\text{Min}_{\zeta} \left[ \text{Max}_{\mathbf{x} \in R} \{ \mathbf{x}^{(m)'} [\mathbf{M}(\zeta)]^{-1} \mathbf{x}^{(m)} \} \right]$$

because  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  is a quadratic form in  $[\mathbf{M}(\zeta)]^{-1}$ . Of course, the natural choices of regions for  $R$  are the cube and sphere. The resulting  $G$ -efficiency is conceptually quite simple



to grasp. It turns out that under the standard independence and homogeneous variance on the model errors, there is a natural lower bound for the maximum prediction variance of

$$\text{Max}_{\mathbf{x} \in R} [v(\mathbf{x})] \geq p \tag{9.14}$$

To understand this bound, consider the expression for hat diagonals discussed in Chapter 2. Given the linear model

$$y_i = \mathbf{x}_i^{(m)'} \mathbf{b} \quad (i = 1, 2, \dots, n)$$

the *i*th hat diagonal is given by

$$h_{ii} = \mathbf{x}_i^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_i^{(m)} \quad (i = 1, 2, \dots, n)$$

and, indeed,

$$\sum_{i=1}^n h_{ii} = \sum_{i=1}^n \mathbf{x}_i^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_i^{(m)} = p \text{ (number of parameters)}$$

See Myers (1990) or Montgomery, Peck, and Vining (2012). Consider a designed experiment and the corresponding scaled prediction variance

$$v(\mathbf{x}) = N \mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)}$$

Suppose we consider the maximum value of  $v(\mathbf{x})$  in some region  $R$ . The *average value* of  $v(\mathbf{x})$  at the data points is  $p$ . Obviously then

$$\text{Max}_{\mathbf{x} \in R} v(\mathbf{x}) \geq p$$

In addition, if a design is  $G$ -optimal, all hat diagonals are equal, thus implying that  $v(\mathbf{x}) = p$  at all hat diagonals and

$$v(\mathbf{x}) \leq p$$

at all other locations in  $R$ . The result in Equation 9.14 is very important and usually very surprising to practitioners. One must keep in mind that  $v(\mathbf{x}) = N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  is a scale-free quantity. The  $G$ -efficiency is very easily determined for a design, say  $\zeta$ , as

$$G_{\text{eff}} = \frac{p}{\text{Max}_{\mathbf{x} \in R} v(\mathbf{x})} \tag{9.15}$$

It is always instructive to investigate efficiencies of our two-level designs of Chapters 4 and 6 with regard to first-order and first-order-plus-interaction models. Consider, for example, a  $2^2$  factorial in the case of the first-order model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

with the region  $R$  being described by  $-1 \leq x_1 \leq 1$  and  $-1 \leq x_2 \leq 1$ . We know that

$$\mathbf{M} = \mathbf{X}'\mathbf{X}/N = \mathbf{I}_3$$

and

$$\begin{aligned} v(\mathbf{x}) &= [1, x_1, x_2]\mathbf{I}_3 \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix} \\ &= 1 + x_1^2 + x_2^2 \end{aligned}$$

As a result, the maximum of  $v(\mathbf{x})$  occurs at the extremes of the cube, namely at  $x_1 = \pm 1$ ,  $x_2 = \pm 1$ . In addition, it is obvious that for the  $2^2$  factorial,

$$\text{Max}_{\mathbf{x} \in R} [v(\mathbf{x})] = 3 = p$$

Thus, the  $G$ -efficiency of the  $2^2$  design is 1.0. As a result, the design is  $G$ -optimal. In fact, it is simple to show that for the first-order model in  $k$  design variables for the cuboidal region, a two-level design of resolution  $\geq \text{III}$  with levels at  $\pm 1$  results in

$$\text{Max}_{\mathbf{x} \in R} [v(\mathbf{x})] = p$$

Thus all these designs are  $G$ -optimal for the first-order model.

Another prediction-oriented optimality is  **$V$ -optimality**. The  $V$ -criterion considers the prediction variance at a specific set of points of interest in the design region, say  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ . The set of points could be the candidate set from which the design was selected, or it could be some other collection of points that have specific meaning to the experimenter. A design that minimizes the average prediction variance over this set of  $m$  points is a  **$V$ -optimal design**.

An important design-optimality criterion that addresses prediction variance is  **$I$ -optimality**, or  $IV$  (integrated variance) optimality. The attempt here is to generate a single measure of prediction performance through an averaging process; that is,  $v(\mathbf{x})$  is averaged over some region of interest  $R$ . The averaging is accomplished via the integration over  $R$ . The corresponding division by the volume of  $R$  produces an average. Thus, the  $I$ -criterion is clearly related to the  $V$ -criterion. The  $I$ -optimal design is given by

$$\text{Min}_{\zeta} \frac{1}{K} \int_R v(\mathbf{x}) d\mathbf{x} = \text{Min}_{\zeta} I(\zeta) \quad (9.16)$$

where  $K = \int_R d\mathbf{x}$ . Notice, then, that we can write the criterion for a design  $\zeta$  as

$$\text{Min}_{\zeta} \left\{ \frac{1}{K} \int_R \mathbf{x}^{(m)'} [\mathbf{M}(\zeta)]^{-1} \mathbf{x}^{(m)} d\mathbf{x} \right\} = \text{Min}_{\zeta} I(\zeta) \quad (9.17)$$

In this way the  $I$ -optimal, or  $IV$ -optimal, design is that in which the average scaled prediction variance is minimized. Strictly speaking, the  $I$ -criterion given in Equation 9.17 is a special case of a more general criterion in which the scaled prediction variance is multiplied by a

weight function  $w(\mathbf{x})$ . The utility of the weight function is to weight more heavily certain parts of  $R$ . Of course, in most practical situations  $w(\mathbf{x}) = 1.0$ .

The  $I$ -optimality criterion is conceptually a very reasonable device to use for choosing experimental designs, since selecting a design with good average prediction variance should produce satisfactory results throughout the design space. The concept in Equation 9.17 in a different form does enjoy some use in computer-generated designs. However, neither  $G$  nor  $I$  is used as much as  $D$ -optimality. We will shed more light on this in the next section. The  $I$ -criterion leads nicely to a  $I$ -efficiency for design  $\zeta^*$  as

$$I_{\text{eff}} = \text{Min}_{\zeta} [I(\zeta)] / I(\zeta^*)$$

As in previous cases, it is instructive to apply the  $I$ -criterion to a simple case, namely a first-order model. Consider, again, the model

$$y = \beta_0 + \sum_{j=1}^k \beta_j x_j + \varepsilon$$

Let us assume that the region is once again given by the cuboidal region  $-1 \leq x_j \leq 1$ . The  $I$ -criterion can be written

$$\begin{aligned} I(\zeta) &= \frac{1}{K} \int_R v(\mathbf{x}) d\mathbf{x} \\ &= \frac{N}{K} \int_R \mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)} d\mathbf{x} \end{aligned} \tag{9.18}$$

where  $\mathbf{x}^{(m)'} = [1, x_1, x_2, \dots, x_k]$ . Here, of course,  $\int_R$  implies  $\int_{-1}^1 \int_{-1}^1 \dots \int_{-1}^1$ . Because  $\mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)}$  is a scalar, we can write

$$\begin{aligned} I(\zeta) &= \frac{N}{K} \int_R \text{tr}[\mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)}] d\mathbf{x} \\ &= \frac{N}{K} \int_R \text{tr}[\mathbf{x}^{(m)} \mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1}] d\mathbf{x} \\ &= \text{tr} \left\{ [\mathbf{X}'\mathbf{X}]^{-1} \left[ \frac{1}{K} \int_R \mathbf{x}^{(m)} \mathbf{x}^{(m)'} d\mathbf{x} \right] \right\} \end{aligned} \tag{9.19}$$

Now, the matrix  $(1/K) \int_R \mathbf{x}^{(m)} \mathbf{x}^{(m)'} d\mathbf{x}$  is a  $(k+1) \times (k+1)$  matrix of **region moments**, that is, the diagonal elements (apart from the initial diagonal element, which is 1.0) are the  $(1/K) \int_R x_i^2 d\mathbf{x}$ . For the off-diagonals, we have  $(1/K) \int_R x_i d\mathbf{x}$  in the first row and first column and  $(1/K) \int_R x_i x_j d\mathbf{x}$  for  $i \neq j$  as the other off-diagonal elements. Because the region  $R$  is symmetric,  $\int_R x_i d\mathbf{x}$  and  $\int_R x_i x_j d\mathbf{x}$  are both zero. As a result, Equation 9.19 becomes

$$I(\zeta) = \text{tr}\{[\mathbf{M}(\zeta)]^{-1} [\mathbf{D}_R]\} \tag{9.20}$$

where  $\mathbf{D}_R$  is a diagonal matrix given by

$$\mathbf{D}_R = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \frac{1}{3}I_k & \\ 0 & & & \end{bmatrix}$$

because  $(1/K) \int_R x_i^2 d\mathbf{x} = (1/2^k)(2^k)(\frac{1}{3}) = \frac{1}{3}$ .

Now, Equation 9.20 can be written as a simple function of the variances of the least squares estimators  $b_0, b_1, b_1, \dots, b_k$  of the coefficients in the first-order model. In fact, we have

$$I(\zeta) = N \left[ \text{Var}(b_0) + \frac{1}{3} \sum_{i=1}^k \text{Var}(b_i) \right] \quad (9.21)$$

The value of  $I(\zeta)$ , the scaled prediction variance, averaged over the cuboidal region, is minimized by a design for which the variances of all coefficients are simultaneously minimized. This, as we learned in Chapter 8, is accomplished by the **variance-optimal** design, namely the two-level first-order orthogonal design (resolution  $\geq$ III) with levels set at the  $\pm 1$  extremes.

If we change the region  $R$  from the cuboidal region to a spherical region, the two-level resolution  $\geq$ III design is both  $D$ - and  $I$ -optimal. Again, the requirement remains that the design be first-order orthogonal with two levels set at the perimeter of the experimental region (see Exercise 9.8).

### 9.2.2 Use of Design Efficiencies for Comparison of Standard Second-Order Designs

Design optimality is an extremely interesting and useful concept. Considerable interest has been drawn to design of experiments and RSM because of the existence of computer packages that allow the user to generate designs with the use of the notion of alphabetic optimality.  $D$ -optimality in particular has received much usage and attention, primarily because it lends itself to relatively simple computation and because algorithms were developed early to implement it. Both Design-Expert and JMP can generate  $D$ -optimal and  $I$ -optimal designs.

The reader should bear in mind the 11 desirable properties of an RSM experimental design, listed in Section 8.1. They suggest that what is required in practice, if they can be obtained, are designs that reflect several important properties, not designs that are optimal with respect to a single criterion. They suggest that what one desires are **good designs**—that is, designs that will produce reliable results under a wide variety of possible circumstances, not designs that are optimal under a fairly restrictive and idealized set of conditions and assumptions. In the case of first-order models and first-order-plus-interaction models, standard designs are optimal designs. We have seen some evidence of this in this chapter; but, in the second-order case, standard RSM designs that we discussed in Chapter 8 are rarely optimal designs, and yet we know that they can be constructed to achieve many desirable properties listed among the 11 given in Chapter 8.

For purposes of an RSM analysis, the standard central composite or Box–Behnken designs should be used whenever possible. When economical designs are needed (near-saturated), the Hoke or hybrid designs should be used whenever possible. *This does not imply that optimal designs obtained through computer-generated designs have no value.* They can be extremely useful and perhaps necessary in many circumstances. There are many conditions encountered by the user under which the use of a CCD or BBD is impossible. In particular, we shall see in Chapter 13 that the use of **mixture designs** often necessitates the use of computer-generated designs due to restrictions on the factor levels and design sizes. Even in nonmixture situations, constraints or unusual design size will often rule out the use of a standard design. We focus on optimal designs in the context of computer-generated designs later in this chapter. But, it seems reasonable that the reader may gain some further insight into the virtues of standard designs if we discuss the performance of these designs in the context of design efficiencies. In what follows, we investigate the CCD, BBD, hybrid, and certain other saturated or near-saturated designs.

It should be emphasized that because the notion of design optimality as introduced by Kiefer considers an experimental design as a probability measure, most finite designs (i.e.,  $N$ -point designs) will naturally have design efficiencies less than 1.0. As a result, it is natural that designs with small design size will have small efficiencies. The information in Table 9.2 give  $D$ - and  $G$ -efficiencies for various CCD, BBD, and hybrid designs for a spherical region. The table includes values of  $k$  from 2 through 5. For the CCD,  $\alpha = \sqrt{k}$ . [Further information regarding these design efficiencies can be found in Lucas (1974, 1976) and in Nalimov, Golikova, and Nikeshina (1970).]

**The central composite and Box–Behnken designs are quite efficient.** For example, note that a CCD for  $k=3$  with three center points has an efficiency of 97.63%. This means that if the design were replicated  $1/0.9763 \cong 1.02$  times, it would be equivalent to a  $D$ -optimal design. Obviously, this design is nearly  $D$ -optimal. Note that while the design efficiencies are highest for  $n_c = 2$ , one should not view these results as an indication that there is no advantage in using additional center runs. The use of three to five center runs, according to our recommendation in Chapter 8, brings stability in  $v(\mathbf{x})$  and provides important pure error degrees of freedom. Note that for  $k=4$ , the CCD and BBD have identical efficiencies. This stems from the fact that the BBD is merely a rotation of the CCD for  $k=4$ . The CCD is rotatable, and the variance properties do not change with rotation of the design.

The lack of  $G$ -efficiency in the CCD for a single center run should not be surprising. Recall that for  $n_c = 0$  and a spherical region,  $\mathbf{X}'\mathbf{X}$  is either singular or near-singular. The use of only one center run results in a large value of  $v(\mathbf{x})$  at the design center. As additional center runs are added, values of  $v(\mathbf{x})$  grow nearer to a condition of stability; it is the nature of the  $G$ - and  $D$ -criteria to react adversely to the use of center runs in most cases. However, in the spirit of good balanced performance across several criteria for good designs, this should not preclude the use of multiple center runs as they serve other important goals.

It is natural to expect saturated or near-saturated designs to have low efficiencies. As a result, the results for some of the hybrid designs are quite impressive, especially in the case of the 311B and 416A, 416B, and 416C. Though we only included results for  $k$  from 2 through 5, it should be noted that the  $D$ -efficiencies and  $G$ -efficiencies for the 628A are both 100%, while the corresponding efficiencies for the 628B with a single center run are 90.6% and 84.4%. Further comparisons of hybrid designs with other economical designs such as the small composite design will be illustrated later. While we will not show tables

**TABLE 9.2 D- and G-Efficiencies for Some Standard Second-Order Designs for a Spherical Region**

| Number of Factors | Sample Size | Design                                 | $D_{\text{eff}}$ (%) | $G_{\text{eff}}$ (%) |
|-------------------|-------------|--|----------------------|----------------------|
| 2                 | 9           | CCD ( $n_c = 1$ )                      | 98.62                | 66.67                |
| 2                 | 10          | CCD ( $n_c = 2$ )                      | 99.64                | 96.0                 |
| 2                 | 11          | CCD ( $n_c = 3$ )                      | 96.91                | 87.27                |
| 3                 | 15          | CCD ( $n_c = 1$ )                      | 99.14                | 66.67                |
| 3                 | 16          | CCD ( $n_c = 2$ )                      | 99.61                | 94.59                |
| 3                 | 17          | CCD ( $n_c = 3$ )                      | 97.63                | 89.03                |
| 3                 | 13          | BBD ( $n_c = 1$ )                      | 97.0                 | 76.92                |
| 3                 | 14          | BBD ( $n_c = 2$ )                      | 96.53                | 71.43                |
| 3                 | 15          | BBD ( $n_c = 3$ )                      | 93.82                | 66.67                |
| 3                 | 10          | 310                                    | 72.8                 | 47.4                 |
| 3                 | 11          | 311A                                   | 79.1                 | 78.6                 |
| 3                 | 11          | 311B                                   | 83.1                 | 90.9                 |
| 4                 | 25          | CCD ( $n_c = 1$ )                      | 99.23                | 60.0                 |
| 4                 | 26          | CCD ( $n_c = 2$ )                      | 99.92                | 98.90                |
| 4                 | 27          | CCD ( $n_c = 3$ )                      | 98.86                | 95.24                |
| 4                 | 25          | BBD ( $n_c = 1$ )                      | 99.23                | 60.0                 |
| 4                 | 26          | BBD ( $n_c = 2$ )                      | 99.92                | 98.90                |
| 4                 | 27          | BBD ( $n_c = 3$ )                      | 98.86                | 95.24                |
| 4                 | 16          | 416A                                   | 81.7                 | 88.0                 |
| 4                 | 17          | 416A ( $n_c = 1$ )                     | 90.6                 | 74.3                 |
| 4                 | 16          | 416B                                   | 96.3                 | 63.1                 |
| 4                 | 17          | 416B ( $n_c = 1$ )                     | 95.1                 | 70.1                 |
| 4                 | 16          | 416C                                   | 96.9                 | 78.0                 |
| 5                 | 43          | CCD ( $n_c = 1$ )                      | 98.6                 | 48.84                |
| 5                 | 44          | CCD ( $n_c = 2$ )                      | 99.60                | 87.63                |
| 5                 | 45          | CCD ( $n_c = 3$ )                      | 99.28                | 85.56                |
| 5                 | 27          | CCD ( $\frac{1}{2}$ rep) ( $n_c = 1$ ) | 98.43                | 77.78                |
| 5                 | 28          | CCD ( $\frac{1}{2}$ rep) ( $n_c = 2$ ) | 98.10                | 87.64                |
| 5                 | 29          | CCD ( $\frac{1}{2}$ rep) ( $n_c = 3$ ) | 96.57                | 84.62                |
| 5                 | 41          | BBD ( $n_c = 1$ )                      | 97.93                | 51.22                |
| 5                 | 42          | BBD ( $n_c = 2$ )                      | 98.83                | 90.91                |
| 5                 | 43          | BBD ( $n_c = 3$ )                      | 98.41                | 88.79                |

of design efficiencies for the case of the cuboidal region, the following represent a synopsis of results for the CCD, Hoke, and Box–Draper designs. Again, for specific construction of the Hoke or Box–Draper design, one should consult Hoke (1974) or Box and Draper (1974).

1. The CCD for  $\alpha = 1.0$  is quite efficient, particularly in the  $D$  sense. For  $k = 3, 4,$  and  $5$ , this design has  $D$ -efficiency from approximately 87% to 97% *without center runs*. Center runs reduce  $D$ -efficiency and change  $G$ -efficiency very little.
2. The Hoke  $D_2$  and  $D_6$  designs have  $D$ -efficiencies that range from 80% to 97%. This is quite good given the small design size. The  $G$ -efficiencies are considerably lower, but this is not unexpected with small designs. In general, the Hoke designs perform better than the Box–Draper designs.

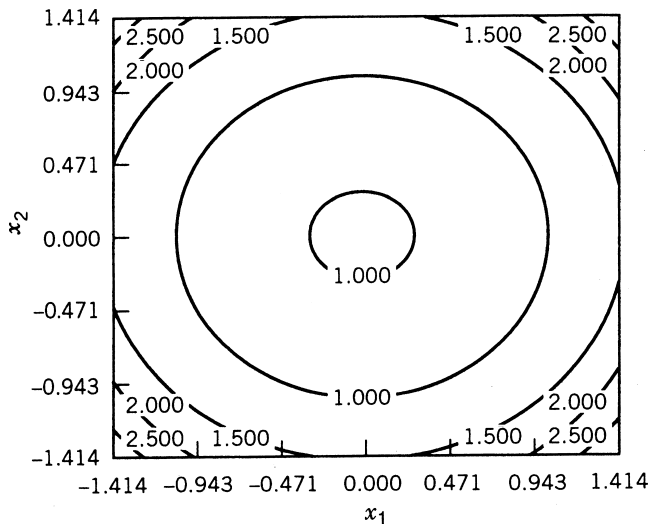
In the next section we focus once again on design comparison. We make use of graphical methods in order to gain some perspective about the distribution of  $v(\mathbf{x})$  in the design region. Our purpose is to compare standard designs and also point the reader toward graphical methodology that can be used to evaluate designs in practice.

**9.2.3 Graphical Procedure for Evaluating the Prediction Capability of an RSM Design**

Design efficiencies are single numbers that purport to evaluate the capabilities of a designed experiment. In cases of interest here, the designs are RSM designs. Design efficiencies can be beneficial. Indeed, they were of benefit to us in the previous section. They enabled the reader to understand that standard RSM designs are often fairly close to the “best possible” design in the  $D$ -optimality or  $G$ -optimality sense. One must remember that in almost all cases (hybrid 628A being an exception) the best finite-sample-size designs have efficiencies less than 100%.

When one is interested in stability of  $v(\mathbf{x})$  over the design region, a single-number criterion such as a design efficiency may not capture the true design capability, which is actually a multidimensional concept. There is useful information in methodology that allows the user to know values of  $v(\mathbf{x})$  throughout the design region. Obviously, for  $k = 2$  two-dimensional plots of  $v(\mathbf{x})$  can easily be constructed. This type of plot can be used to compare designs. In addition, for a specific RSM analysis, if data have been collected, a plot showing contours of constant  $\widehat{\text{Var}}[\hat{y}(\mathbf{x})]$  or  $\{\widehat{\text{Var}}[\hat{y}(\mathbf{x})]\}^{1/2}$  can aid the practitioner in visualizing confidence limits on the mean response at interesting locations in the design region. The following example provides an illustration.

**Example 9.1 Contour Plots of Prediction Variance** Consider the  $k = 2$  illustration of Example 6.4. The design used is a second-order rotatable CCD. Contours of constant response are shown in Fig. 6.16. Figure 9.3 is a two-dimensional plot showing the estimated



**Figure 9.3** Contours of constant standard error of prediction.

prediction standard deviation or estimated standard error of prediction

$$\{\text{Var}[\widehat{y}(\mathbf{x})]\}^{1/2} = s\sqrt{\mathbf{x}^{(m)'}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}^{(m)}}$$

where  $s$  is the square root of the mean square error. Note the stability of the prediction standard error on spheres. Note also that the prediction begins to break down as one moves outside the design region. For example, for  $\mathbf{x}_1 = \mathbf{x}_2 = 1.414$  at the corner of the plot, we are outside the design region and beyond where any data have been observed.

For values of  $k > 2$ , contours of prediction variance are, of course, considerably more awkward. However, there are other graphical procedures that allow comparison of the distribution of  $v(\mathbf{x})$  among RSM designs for any number of design variables. These graphical procedures make use of **variance dispersion graphs** and **fraction of design space plots**. The following sub-subsections define both types of plots and illustrate their use.

**Variance Dispersion Graphs (VDGs)** Giovannitti-Jensen and Myers (1989) and Myers et al. (1992b) developed the variance dispersion graph and produced case studies that allowed for interesting comparisons between standard designs. Rozum (1990) and Rozum and Myers (1991) extended the work from spherical to cuboidal regions.

A variance dispersion graph for an RSM design displays a “snapshot” that allows the practitioner to gain an impression regarding the stability of  $v(\mathbf{x}) = N \text{Var}[\widehat{y}(\mathbf{x})]/\sigma^2$  and how the design compares to an ideal. Let us first consider a spherical design region. The VDG contains three graphical components:

1. A line of the **spherical variance**  $V^r$  against the radius  $r$ . The spherical variance is essentially  $v(\mathbf{x})$  averaged *over the surface* of a sphere of radius  $r$ , namely,

$$V^r = \frac{N\psi}{\sigma^2} \int_{U_r} \text{Var}[\widehat{y}(\mathbf{x})] d\mathbf{x} \quad (9.22)$$

where  $U_r$  implies integration over the surface of a sphere of radius  $r$  and  $\psi = (\int_{U_r} d\mathbf{x})^{-1}$ .

2. A line of the *maximum*  $v(\mathbf{x})$  on a radius  $r$  against  $r$ . That is,

$$\max_{\mathbf{x} \in U_r} \left[ \frac{N \text{Var}[\widehat{y}(\mathbf{x})]}{\sigma^2} \right] = \max_{\mathbf{x} \in U_r} [v(\mathbf{x})]$$

3. A line of the *minimum*  $v(\mathbf{x})$  on a radius  $r$  against  $r$ .

In addition to the above three lines, which offer information about  $v(\mathbf{x})$  for the design being evaluated, the result of Equation 9.14 is used to provide some sense of closeness to ideal. A horizontal line at  $v(\mathbf{x}) = p$  is added. Obviously, if the design is rotatable, the entire VDG will involve only the spherical variance and the horizontal line, since the maximum, minimum, and average  $v(\mathbf{x})$  will be identical at any radius  $r$ .



It is of interest to discuss briefly the computations associated with the spherical variance. It turns out that the integral in Equation 9.22 can be written as

$$\begin{aligned}
 V^r &= \frac{N\psi}{\sigma^2} \int_{v_r} \text{Var}[\hat{y}(\mathbf{x})] d\mathbf{x} \\
 &= \frac{N\psi}{\sigma^2} \int_{v_r} \mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)} d\mathbf{x} \\
 &= \frac{N\psi}{\sigma^2} \int_{v_r} \text{tr}[\mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)}] d\mathbf{x} \\
 &= \frac{N\psi}{\sigma^2} \int_{v_r} \text{tr}[(\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)} \mathbf{x}^{(m)'}] d\mathbf{x} \\
 &= \text{tr}(\mathbf{X}'\mathbf{X})^{-1} \left[ \frac{N\psi}{\sigma^2} \int_{v_r} \mathbf{x}^{(m)} \mathbf{x}^{(m)'} d\mathbf{x} \right] \\
 &= \text{tr}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{S}
 \end{aligned} \tag{9.23}$$

where  $\mathbf{S}$  is the  $p \times p$  matrix of region moments, with the region being the hypersphere given by  $U_r$ .

**Fraction of Design Space (FDS) Plots** Zahran et al. (2003) developed the fraction of design space plots as a complement to the variance dispersion graph. While the variance dispersion plot shows where in the design space various values of  $v(\mathbf{x})$  are observed, the fraction of design space plots summarizes the prediction performance for the entire design space with a single curve. The stability of  $v(\mathbf{x})$ , its minimum, maximum, and quantiles can all be extracted from the summary curve, allowing for easy assessment of the prediction of a design or comparisons between competing designs. The FDS for an ideal design will have a large fraction of the design space with small  $v(\mathbf{x})$  values and be relatively flat, which corresponds to stability of  $v(\mathbf{x})$  throughout the region.

The curve of an FDS plots matches the value of  $v(\mathbf{x})$  against the fraction of the design space that has prediction variance less than or equal to the given value. A graph of the cumulative fraction provides the user with a sketch of the prediction variance throughout the appropriate design region. The plot can be constructed by taking a random sample of design locations throughout the design space, and calculating  $v(\mathbf{x})$  at these locations. These values are then sorted and plotted against the quantiles from 0 to 1, in a manner similar to a cumulative distribution function for a statistical distribution. Both JMP and Design-Expert can construct FDS plots for a given design.

By using VDGs and FDS plots in combination, we can make comparisons between different potential designs before any data are collected. It is important when we compare competing designs to make sure that the same design regions are being compared, and that these represent the area of interest for the experiment. The FDS plot allows for direct comparisons of which design has a larger proportion of the design space at or below a particular value, or for a fixed fraction, which design has the smaller  $v(\mathbf{x})$  value. Each design is summarized with a single curve, which helps make comparisons straightforward. The maximum  $v(\mathbf{x})$  value is shown in both VDG and FDS plots allowing for easy observation of the  $G$ -efficiency. The VDGs allow us to see where the good and bad regions of prediction

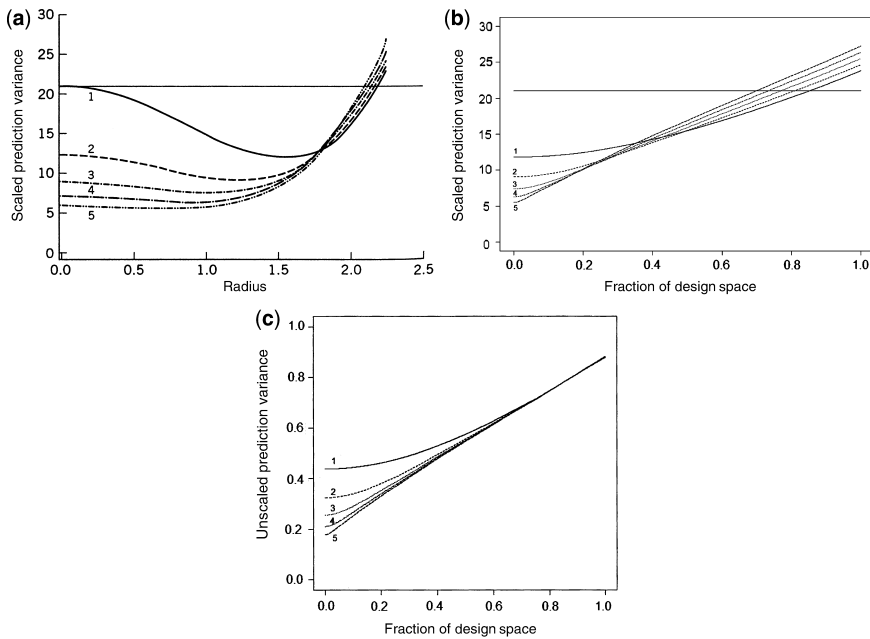
are relative to the distance from the center of the design, as well as assess if the design is rotatable or near-rotatable. One disadvantage of the VDGs is that the relative emphasis of different regions of the design space are not proportionate to the fraction of the design space that they represent. Small values of the radius (namely those close to zero) where only a small proportion of the design space is located are given greater visual emphasis than the edge of the region. This effect becomes increasingly dramatic as the number of factors,  $k$ , increases.

The goal of using graphical summaries to compare different designs is to gain better understanding of where the designs predict well or poorly in the design space of interest, as well as to see how well we might expect to predict future observations based on the experiment that we will have run.

In the following example we illustrate the use of the VDG and FDS plots in a study to determine the utility of center runs with a rotatable CCD for  $k = 5$ .

**Example 9.2 Choosing the Number of Center Runs in a CCD** Consider a central composite design in five design variables with the factorial portion being the resolution V fraction of a  $2^5$ . The value of  $\alpha$  is 2.0, so that the design is rotatable. As a result,  $v(\mathbf{x})$  is constant on spheres, and hence  $\max_{\mathbf{x} \in R} [v(\mathbf{x})] = \min_{\mathbf{x} \in R} [v(\mathbf{x})] = V'$ . The plot in Fig. 9.4a shows five VDGs for the number of center runs  $n_c = 1$  through  $n_c = 5$ . The horizontal line is at  $p = 21$ . Recall that the maximum value of  $v(\mathbf{x})$  in the design region can be no smaller than  $p$ , and if this value is attained, we have a  $G$ -optimal design. Figure 9.4b shows the corresponding FDS plot for the same five designs.

The plot in Fig. 9.4a reflects values in the region  $0 \leq r \leq \sqrt{5}$  because the factorial points are at radius  $\sqrt{5}$ . The VDG shows how additional center runs lower  $v(\mathbf{x})$  at the center of

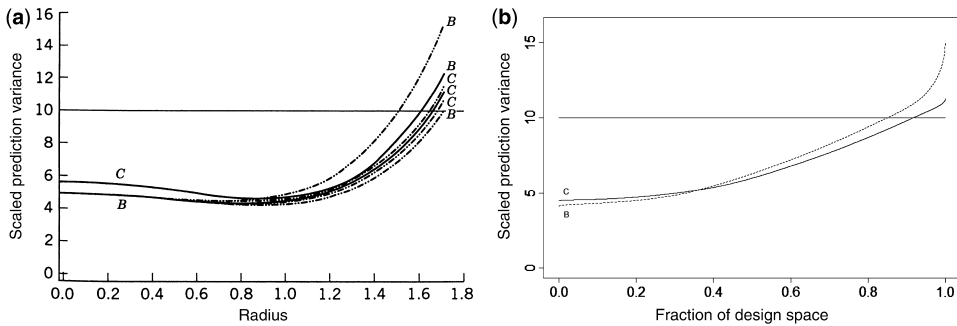


**Figure 9.4** Plots showing the effect of center runs for a  $k = 5$  rotatable CCD with a  $2^{5-1}$  as the factorial portion (a) VDG (b) FDS plot with SPV, (c) FDS plot with UPV.

the design space at the cost of slightly worsening prediction at the edge of the design space. This worsening is a function of  $v(\mathbf{x})$  penalizing larger designs. The rate of improvement at the center is diminishing and becomes marginal for additional center runs after  $n_c = 4$ . The FDS shows the same information in a different way, with the minimum  $v(\mathbf{x})$  shown at the left of the plot, and the maximum  $v(\mathbf{x})$  at the right for each design. In terms of  $G$ -efficiency, the best design is the CCD with  $n_c = 1$ , but this has the largest minimum  $v(\mathbf{x})$  value. Notice how the VDG accentuates the improvement at the center of the design space, even though this is proportionately a much smaller fraction of the design region than at the edge of the design space.

In Fig. 9.4c, a FDS plot is shown for the unscaled prediction variance. As discussed in Chapter 8, the unscaled prediction variance is useful to consider when we are interested in the raw improvement in prediction variance for additional runs of our experiment. In this plot, we see the effect of the additional center runs on the overall prediction performance. With the scaled prediction variance, there was a trade-off in prediction performance on a per observation basis, with improvements at the center of the design space resulting in a slightly worse  $v(\mathbf{x})$  at the edge of the design space. With the unscaled prediction variance, we see that the additional center runs (and corresponding increase in design size) can only improve the prediction. There are slight gains throughout a majority of the design space, but the maximum unscaled prediction variance at the edge of the design space is unaffected by additional observations at the center of the design space.

**Example 9.3 Comparison of the CCD and the BBD** In this illustration we use the variance dispersion graph and fraction of design space plot to show a comparison of the distribution of  $v(\mathbf{x})$  in the design region between a CCD and a BBD. The reader should recall that both of these very popular designs are spherical designs. Both designs involve three design variables, with the CCD having  $\alpha = \sqrt{3}$  and  $n_c = 3$ , while the BBD also contains three center runs. One must keep in mind that the total sample sizes are 17 and 15 for the CCD and BBD, respectively. A plot with both VDGs is in Fig. 9.5a. Note that both designs contain three lines in the VDG; the min  $v(\mathbf{x})$ , average  $v(\mathbf{x})$ , and max  $v(\mathbf{x})$  are shown because neither design is rotatable. The plots stop at radius  $\sqrt{3}$ . Both designs have been scaled so that points are at a radius  $\sqrt{3}$  from the design center. Figure 9.5b shows the FDS plot for the same two designs.



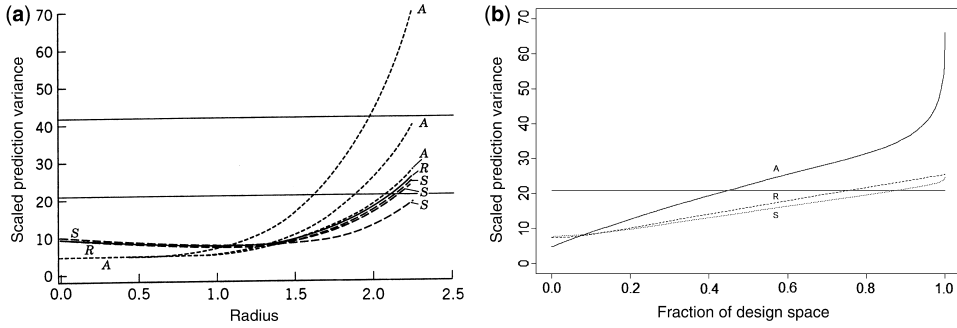
**Figure 9.5** Variance dispersion graphs (a) and fraction of design space plots (b) for CCD (C) and Box-Behnken design (B) for  $k = 3$ . Both designs contain three center runs.

The following are observations and conclusions from the plots.

1. The combination of the VDG and FDS plots provide a wealth of information about the two designs being compared. From the FDS plot it is easy to see that the BBD has better minimum prediction than the CCD, but at the cost of worse maximum prediction at the edge of the design space. The minimum  $v(\mathbf{x})$  for both designs occurs at a radius between 0.8 and 1.0.
2. The values of  $v(\mathbf{x})$  become much larger for the BBD at the very edge of the region, but only a very small proportion of the region has these very large  $v(\mathbf{x})$  values. Hence  $G$ -efficiency, which considers only the maximum  $v(\mathbf{x})$  can be somewhat deceptive about what is happening with the majority of the design space.
3. Comparison with the ideal design is readily seen for both designs. The maximum  $v(\mathbf{x})$  in the design region is roughly 15 for the BBD and 11.2 for the CCD. Thus the  $G$ -efficiencies for the two designs are  $(10/11.2) \times 100 = 89\%$  for the CCD and  $(10/15) \times 100 = 67\%$  for the BBD. These numbers correspond to those given in Table 9.2.
4. Comparing the two designs is somewhat easier using the FDS plot since each design is summarized with a single curve, while the non-rotatability of the designs makes comparing using the three curves with the VDG a bit more difficult. However, being able to see where the good and bad  $v(\mathbf{x})$  values occur is only available using the VDGs.
5. The  $\alpha = \sqrt{3}$  CCD is nearly rotatable. This is not surprising because  $\alpha = 1.682$  results in exact rotatability. The slight deviation from rotatability occurs close to the design perimeter. The BBD shows a more pronounced deviation from rotatability that becomes greater near the edge of the design space.
6. The values of  $v(\mathbf{x})$  are very close for the two designs near the design center. In fact the difference at the center is accounted for by the difference in the sample sizes.
7. For prediction from radius 1.0 to  $\sqrt{3}$  the CCD is the better design. The stability in  $v(\mathbf{x})$  is greater with the CCD, and the max  $v(\mathbf{x})$  is smaller than that of the BBD. This is also reflected in the FDS plot where for all fractions along the  $x$ -axis greater than 0.4, the CCD has the smaller  $v(\mathbf{x})$  value.

The illustration in this example nicely underscores the usefulness of graphical methods for comparing competing designs. It also illustrates how a single-number efficiency, while useful, cannot capture design performance. For example, the clear distinction between the two designs in terms of prediction errors is not captured in the comparison of  $D$ -efficiencies (97.63% and 93.82%) from Table 9.2. In addition, the  $G$ -efficiencies accurately reflect the differences that exist in  $v(\mathbf{x})$ , *but only at the design perimeter*. Almost without exception (completely without exception in the first-order case), max  $v(\mathbf{x})$  in a design region for a second-order design occurs at the design perimeter. Thus  $G$ -efficiency usually reflects what is occurring at the design perimeter.

**Example 9.4 Comparison of Different Axial Values of the CCD** In Fig. 9.6 we show the graphical summaries for three CCDs with  $k=5$  in a spherical region with maximum radius  $\sqrt{5}$ . The  $\alpha$ -values compared are 1.5, 2.0, and  $\sqrt{5}$ . In each case the factorial portion



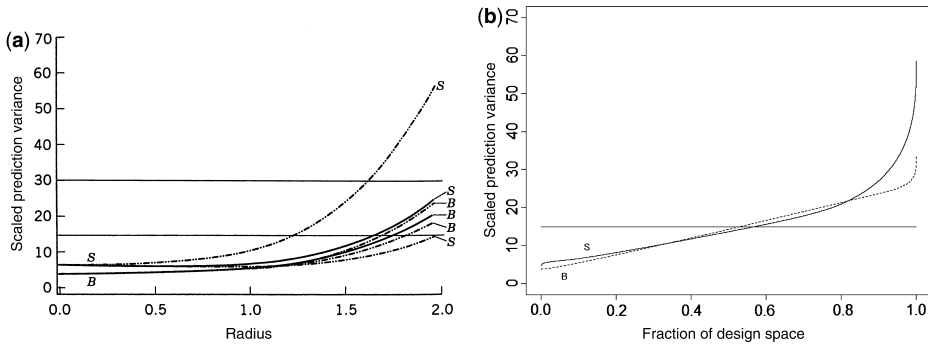
**Figure 9.6** VDG and FDS for three CCDs for  $k=5$ . Design A sets  $\alpha = 1.5$ , design R sets  $\alpha = 2.0$ , and design S sets  $\alpha = \sqrt{5}$ . Each contains a  $2^{5-1}$  and three center runs.

is a resolution V fraction. Each design contains three center runs, and thus the total design size is  $N=29$ . The  $\alpha = 1.5$  design is not expected to predict as well near the edge of the design space, because the axial points are not placed at or near the region perimeter. The  $\alpha = 2.0$  design is rotatable because  $\alpha = \sqrt[4]{16} = \sqrt[4]{F} = 2.0$ .

The following are conclusions drawn from Fig. 9.6.

1. The use of  $\alpha = 1.5$  does result in a loss in efficiency compared to the other two designs. While the prediction is better in the design center, it suffers at the perimeter. The FDS plot shows that this benefit occurs for less than 10% of the total design space. In fact, one gains some insight into how much efficiency is lost with  $\alpha = 1.5$  by observing that the value of  $\max v(\mathbf{x})$  at the design perimeter is roughly 65. The two horizontal lines in the VDG are at  $p$  and  $2p$ , representing  $G$ -efficiencies of 100% and 50%. The  $G$ -efficiency for the  $\alpha = 1.5$  design is roughly  $(21/65) \times 100 \cong 32\%$ .
2. While  $\alpha = 2.0$  is the rotatable value, the use of  $\alpha = \sqrt{5}$  and the resulting loss of rotatability do not produce severe instability in  $v(\mathbf{x})$ . The same appeared to be true for the  $\alpha = \sqrt{3}$  CCD in Fig. 9.5. Again, this VDG reflects the fact that for the CCD, allowing the value of  $\alpha$  to be pushed to the edge of the design perimeter, namely to  $\sqrt{k}$ , is preferable, even if rotatability is compromised. Clearly the loss in rotatability is a trivial matter here and for the CCD in Fig. 9.5. The  $\max v(\mathbf{x})$  plot for design S or design R reflects a  $G$ -efficiency of roughly 85%, which is consistent with the results in Table 9.2.
3. The choice of  $\alpha = \sqrt{5}$  improves estimation not only at the edge of the design space, but also throughout most of the region. By comparing the lines of the FDS plot, we can see that the lower line for the  $\alpha = \sqrt{5}$  CCD represents a helpful improvement in reducing  $v(\mathbf{x})$ .

One pragmatic point needs to be made here regarding design S in Fig. 9.6. For  $k=5$  and other large values, the “best” CCD requires  $\alpha = \sqrt{k}$ . This requires the levels set by the investigator to be  $-2.236, -1, 0, 1, 2.236$ . This design may, in fact, produce practical problems because the axial level is much larger than the factorial level. Indeed, the evenly spaced levels of design R may be more reasonable. For  $k=2, 3$ , and 4, this is not a serious problem.

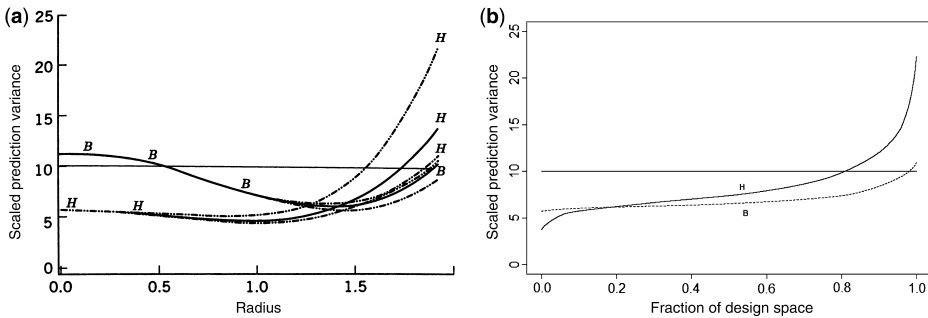


**Figure 9.7** Variance dispersion graphs and fraction of design space plots for hybrid 416B and SCD (S). Design B is hybrid 416B.  $N = 19$  for the SCD and  $N = 20$  for 416B.

**Example 9.5 Hybrid Designs and SCDs** In previous sections we mentioned the importance of the class of hybrid designs in cases where the researcher requires the use of a small design size. The efficiencies of the hybrid design in Table 9.2 are quite good for saturated or near-saturated designs, though this may not be apparent, because we showed no comparisons with other designs of small sample size. In Fig. 9.7a we show overlaid VDGs of a hybrid 416B with four center runs and a  $k = 4$  SCD with three center runs. The small composite contains a resolution III  $2^{4-1}$  with axial value  $\alpha = 2.0$ . Figure 9.7b shows the corresponding FDS plot for the two designs.

From the  $p$  and  $2p$  lines it is clear that the 416B is considerably more efficient. The  $G$ -efficiency of the 416B is roughly 60%, while the  $G$ -efficiency for the SCD is only slightly above 25%. More importantly, the max  $v(\mathbf{x})$  for the SCD becomes large very quickly. The discrepancy between the designs is not only at the design perimeter. The FDS plot shows that the 416B design has a smaller minimum  $v(\mathbf{x})$ , is competitive with the SCD for most fraction of the design, and far superior with a much smaller maximum  $v(\mathbf{x})$ . As the number of factors increases, the maximum  $v(\mathbf{x})$  of the SCD rises very rapidly making this a poor choice of design for good prediction.

**Example 9.6 Adding Center Runs to Hybrid Designs** Figure 9.8 gives overlaid VDGs for the hybrid 310 with a center run added and for the hybrid 311B. One should note from the efficiency table (Table 9.2) that both the  $D$ - and  $G$ -efficiency for the 311B are quite good



**Figure 9.8** VDGs and FDS for 310 plus one center run and for 311B. Design B is hybrid 311B, and design H is hybrid 310.

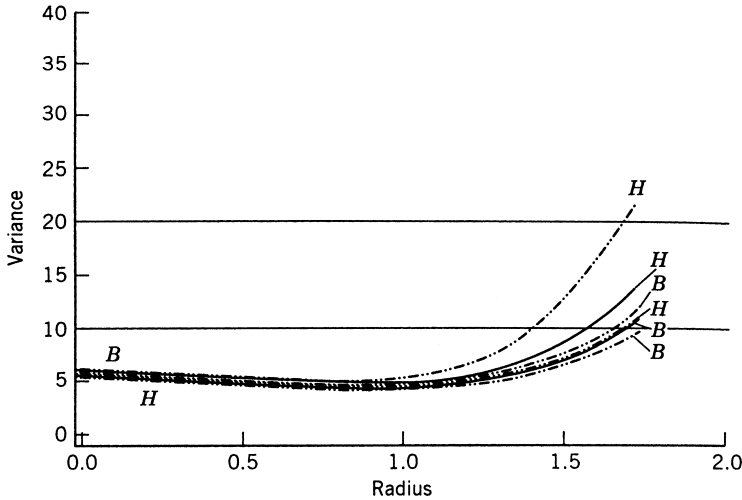


Figure 9.9 VDGs for 310 plus one center run and for 311B plus one center run.

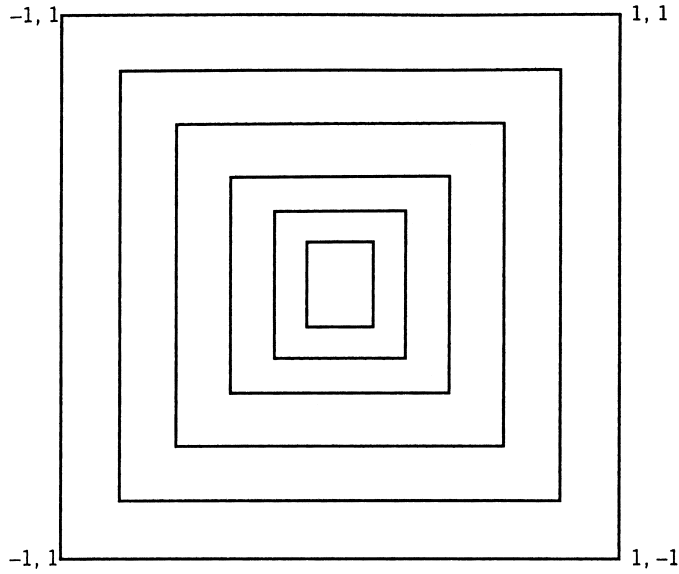
compared to the 310. Indeed, from Fig. 9.8a it is clear that the 311B is nearly rotatable with high  $G$ -efficiency. However, what the efficiency figures do not reflect is that the prediction variance is much smaller for the 310 near the design center. The FDS plot of Fig. 9.8b helps clarify the relative proportion of the design space affected by the poorer prediction at the center by 311B. The apparent advantage of better prediction by 310 represents only a small fraction of the total volume of the design space. The better prediction of the 311B at the perimeter is much more dominant than is highlighted by the VDG. Again, the combination of the VDG and FDS plot allows for improved understanding of  $v(\mathbf{x})$  throughout the design region.

Incidentally, the plots suggests that an additional center run for the 311B would substantially reduce  $v(\mathbf{x})$  at the center. This, in fact, is the case. An additional run in the center reduces  $v(\mathbf{x})$  to approximately the value achieved by the 310. Interestingly, this increases  $v(\mathbf{x})$  at the perimeter slightly, *thereby reducing the  $G$ -efficiency*. Thus, we have often seen situations of a design (311B + one center run) that has a smaller  $G$ -efficiency (and  $D$ -efficiency in this case) than does a clearly inferior design (311B). Figure 9.9 shows a VDG to compare the same two designs as in Fig. 9.8 with the additional center run given to the 311B.

**VDGs for Cuboidal Designs** All of the preceding discussion and illustrations deal with variance dispersion graphs for spherical regions. In this case it is natural to observe values of  $v(\mathbf{x})$  averaging over the surfaces of spheres. Then the minimum and maximum of  $v(\mathbf{x})$  is taken over the surfaces of spheres. However, it is not natural to deal with surfaces of spheres when the design is cuboidal, and hence the natural region of interest is a cube. As a result, the cuboidal variance dispersion graph contains the following components:

1. A plot of the **cuboidal variance**  $V^{r*}$  against  $r$ , the half edge of a cube:  $V^{r*}$  is defined as

$$V^{r*} = \frac{N\psi^*}{\sigma^2} \int_{U_r^*} \left[ \frac{\text{Var}[\hat{y}(\mathbf{x})]}{\sigma^2} \right] d\mathbf{x}$$



**Figure 9.10** Nested cubes of the VDG for cuboidal designs.

where  $U_r^*$  is the surface area of the cube with half edge  $r$  and  $\int_{U_r^*}$  implies integration over the surface area of the same cube. In our illustrations  $r$  is called the **radius**. We define  $\psi^* = (\int_{U_r^*} d\mathbf{x})^{-1}$ .

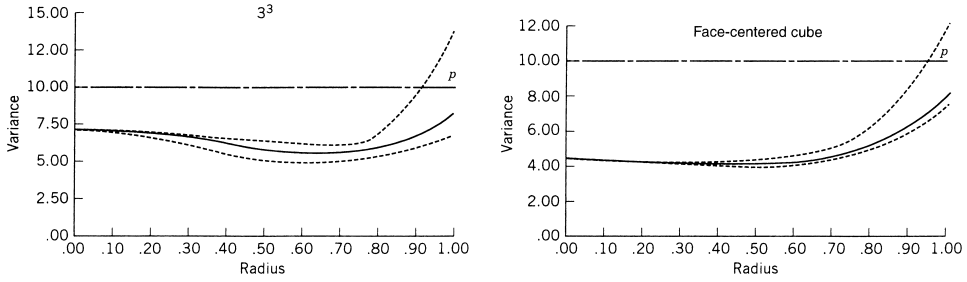
2.  $\max_{\mathbf{x} \in U_r^*} [v(\mathbf{x})]$ .
3.  $\min_{\mathbf{x} \in U_r^*} [v(\mathbf{x})]$ .

As a result, the average and range of  $v(\mathbf{x})$  are depicted over the surface of cubes that are nested inside the cube defined by the design region. Figure 9.10 gives the reader an impression of what is displayed by the VDG. Averages, maxima, and minima are computed for half edges from 0 to 1.0. These values are plotted in a two-dimensional graph. Spheres nested inside the design cube can be used as the media for the graphs. See Myers et al. (1992b). However, cubes nested inside the design cube are more natural, just as rotatability [i.e., constant  $v(\mathbf{x})$  on spheres] is a more important criterion for spherical regions of interest than for cuboidal regions.

**FDS Plots for Cuboidal Regions** The adaption of the FDS plot needed for cuboidal regions is much simpler than for the VDGs. Indeed for any shaped design region, the FDS plot remains unchanged. The  $x$ -axis still is from 0 to 1, and the interpretation of the quantiles is the same. For calculating the values of  $v(\mathbf{x})$ , we sample uniformly from our newly shaped region, calculate  $v(\mathbf{x})$ , and plot the sorted values against the appropriate quantiles.

**Comparison of  $3^3$  Factorial with Face-Centered Cube** Although the CCD involves considerably fewer design points than the three-level factorial design for  $k = 3$ , they are natural rival designs when the design region is a cube. Of course, the scaled prediction variance will involve a sample size weighting of  $N = 27$  for the factorial design and  $N = 15$  for the  $\alpha = 1$  CCD, that is the face-centered cube (FCC). Figure 9.11 shows the VDGs for the two

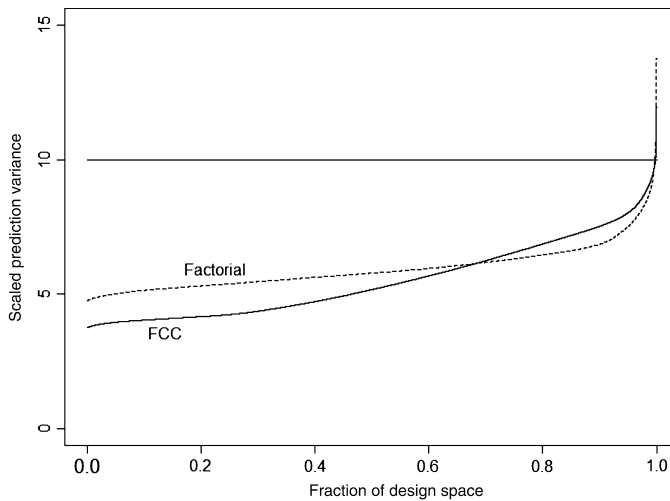




**Figure 9.11** Variance dispersion graphs for the  $3^3$  factorial and the face-centered cube (FCC).

designs, and Fig. 9.12 shows the corresponding FDS plot. The FCC design predicts considerably better near the center of the design space, and has very stable prediction for points of radius 0.6 or less. From the FDS plot we can see that the  $3^3$  factorial design is being heavily penalized for its large design size with poorer performance at both the minimum and maximum  $v(\mathbf{x})$  values. The stability of prediction for the  $3^3$  factorial design is quite good for most of the region, with only a small fraction of the region (less than 2%) having much larger  $v(\mathbf{x})$  values. The  $G$ -efficiency of the FCC is better than that of the  $3^3$  factorial design. Certainly the face-centered cube is an important design for cuboidal regions. However, if cost-considerations allow for a design of size  $N = 27$ , the full  $3^3$  factorial design is a good alternative.

Variance dispersion graphs for cuboidal regions are no less important than those for spherical regions. Both are useful tools for comparing competing designs. The VDG and FDS plot will surface again later in this chapter as we focus on computer-generated designs. Information regarding software for construction of VDGs can be found in Vining (1993), Rozum and Myers (1991), and Borror (1998).



**Figure 9.12** Fraction of design space plot for comparing  $3^3$  factorial and the face-centered cube (FCC).

**Extensions and Alternatives to the VDG and FDS Plots** There have been a number of very useful extensions of the VDG. Trinca and Gilmour (1998) extend the VDG to situations where the response surface design has been run in **blocks**. They point out that allocation of runs to the blocks can have serious effects on the prediction properties of a design. Vining and Myers (1991) extend the VDG to plots of the **average mean square error**, which is the sum of prediction variance and squared bias (use of mean square error and bias in design selection is discussed in Chapter 10). Borror, Montgomery, and Myers (2002) construct VDGs for experiments containing **noise variables**. These designs are used extensively in process robustness studies (see Chapter 11). Vining, Cornell, and Myers (1993) show how to develop VDGs for mixture designs (see Chapters 12 and 13 for an introduction to experiments with mixtures).

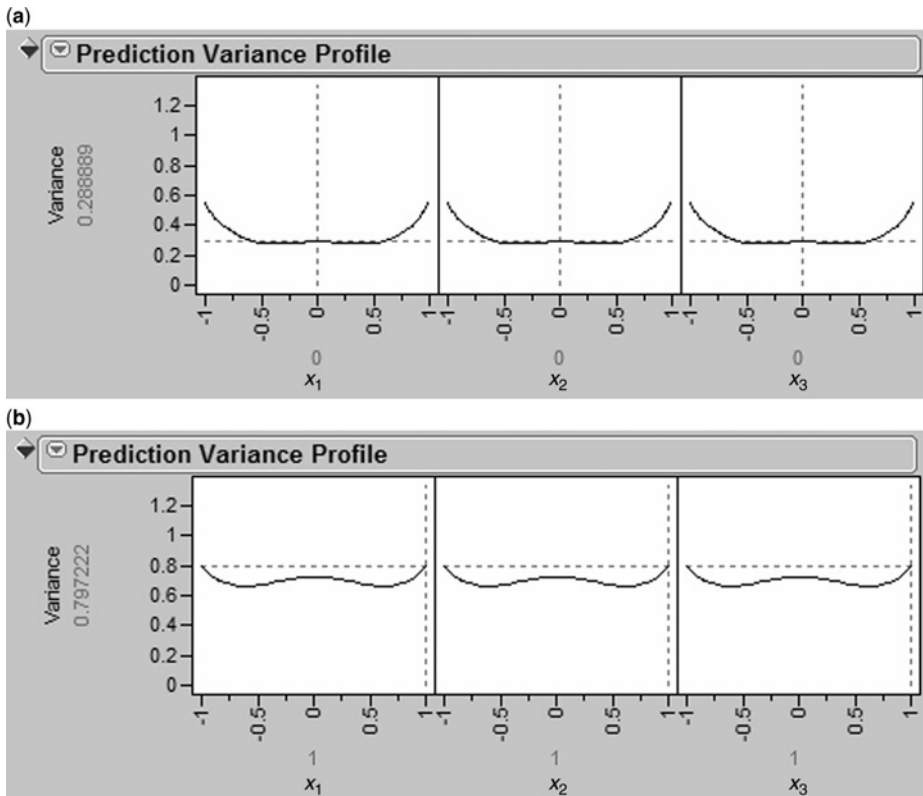
Goldfarb et al. (2004a) developed three-dimensional VDGs for mixture-process experiments, where interest lies in understanding differences in  $v(\mathbf{x})$  separately for the mixture and process variables. Goldfarb et al. (2004b) also adapted FDS plots for mixture and mixture-process experiments. Liang et al. (2006a, 2006b) created VDGs and FDS plots for split-plot designs, where the whole plot and split-plot characteristics could be examined separately. Ozol-Godfrey et al. (2005) and Anderson-Cook et al. (2009) adapted FDS plots to predict when the assumed model may not be correct. Liang et al. (2007) consider alternative FDS plots which adjust for different cost structures for split-plot designs, which allow for flexible comparison of quality of prediction versus the cost of running the experiment.

Khuri et al. (1996) describe a method for describing the entire distribution of prediction variance over the design region. They do this by plotting **quantiles** of the prediction variance distribution. They note that quantile plots give more information than VDGs and present an example where the VDGs of the competing designs are quite similar yet the quantile plots of prediction variance are very different. Khuri, Harrison, and Cornell (1999) extend this technique to experiments with mixtures.

A dynamic graphical tool for visualizing the prediction variance is available in JMP. This allows the user to interactively examine the unscaled prediction variance of a design. Figure 9.13 shows the prediction variance for the face-centered cube for  $k=3$  at both the center (0,0,0) and corner (1,1,1) of the design region. When used in JMP, any combination of factor levels in the design region can be explored. Note that the maximum unscaled prediction variance value of 0.797222 obtained at any corner of the design space matches with the maximum values of  $0.797222 \times 15 = 11.958$  in the VDG and FDS plots of Figs. 9.11 and 9.12.

### 9.3 COMPUTER-GENERATED DESIGNS IN RSM

As we described in Section 9.2, computer-generated designs can provide a flexible complement to classic response surface designs like the Central Composite, Box–Behnken, and Definitive Screening designs. Whereas the classic designs have fixed structure, shapes of design regions, and design sizes, computer-generated designs can be adapted and tailored to different design regions, different numbers of runs, and different priorities. The choice between whether to use a classic response surface design and a computer-generated design for a particular study is not always obvious. The CCD, BBD, and DSD designs have good performance for estimation of model parameters and prediction of new observations. They also often have some other built-in properties of the ability to estimate lack of fit and pure error, which adds robustness to their overall utility. However, if the design size,



**Figure 9.13** Prediction profiler from JMP to show the prediction variance for the face-centered cube for  $k=3$  at the center (a) and corner (b) of the design region.

assumed model, or experimental region for the classical designs do not match well with study requirements, then a computer-generated design can provide useful alternatives.

Computer-generated design of experiments was the natural extension of design optimality as the computer era in statistics unfolded in the late 1960s and early 1970s. Algorithms exist to efficiently find a best design for a specified situation.

Today many computer packages generate designs for the user. Historically, the focus has been on a single criterion, and the criterion most often used is  $D$ -optimality or rather,  $D$ -efficiency. In addition, some computer packages offer a criterion that involves the use of  $y(x)$ , the scaled prediction variance. Common choices are to focus on either  $I$ - or  $G$ -efficiency. The computer is used to aid the researcher in designing RSM experiments in several ways. Some of these are as follows:

1. To design an experiment from scratch.
2. To augment or repair a current designed experiment.
3. To replace points in a current design.
4. To construct an "ideal" standard experimental design for a particular region of interest or to satisfy certain constraints.

There were many important papers that provided the fundamental theory that led to the current state of technology in computer-generated design. The work of Dykstra (1966),

Covey-Crump and Silvey (1970), and Hebble and Mitchell (1972) represent only a few important contributions. Initially, we consider some options if the user chooses to focus on just a single criterion, such as  $D$ -,  $I$ -, or  $G$ -efficiency. These criteria focus on good estimation and prediction under the assumption that the specified model is adequate for the underlying relationship between the input variables and the response, and they reflect the most common focus for design construction based on good design properties described in Section 8.1.

There is a recent movement to formally consider multiple objectives when creating a computer-generated design. Similar to the multiple response optimization choices described in Chapter 7, the common approaches for considering several objectives for a good design involve either (a) conditionally optimizing one criterion subject to satisfying some minimal threshold for performance of another or (b) using Pareto front optimization. Some of these choices are described in Section 9.4.

Typically, the user is required to specify a number of items before a single criterion optimal computer-generated design can be created:

- (i) An assumed model, where some software packages (such as JMP) allow for a designation of model terms as “necessary” or “if possible”, to provide flexibility in the set of models under consideration.
- (ii) The required sample size, where some number of runs may be designated to be replicates.
- (iii) Ranges on the design variables.
- (iv) A set of candidate points.
- (v) Optionally, other possible constraints (e.g., blocking information).
- (vi) The criterion to be optimized.

We note several important aspects about the items that the user must specify. First, perhaps obviously, the computer-generated design will be a direct reflection of these choices. Hence careful thought should be given to ensure that the design is an appropriate reflection of the intended study. Specifying an appropriate model has direct consequences for which locations will be included in the design. Second, the selected sample size can play a substantial role in the quality of estimation and prediction. The user should aim to include sufficient runs to avoid saturated designs whenever possible, as this typically leads to improved design characteristics, and offers some protection if there are problems with the implementation of the design. Next, with a computer-generated design, the software can often incorporate the natural ranges of the variables, while for classic designs the user typically does the translation between the coded and natural variables. One should notice the reference to a list of candidate points. The implication is that the computer picks from this list for design construction or augmentation. If the candidate list is not supplied, often the computer package will use a default grid of points of its choice. Thus, the design chosen certainly is a discrete design, and the goal is to select the design with the highest efficiency, say  $D$ -efficiency, for the specified criterion. We emphasize the importance of understanding and specifically choosing the criterion over which the optimization will be performed. An “optimal design” having used an undesirable or irrelevant criterion will not necessarily provide a useful solution for the user. With increased computer power, finding optimal designs has become more tractable and timely.

### 9.3.1 Important Relationship Between Prediction Variance and Design Augmentation for $D$ -Optimality

An important relationship exists between the coordinates of points that augment an existing design via  $D$ -optimality and the concept of prediction variance. In fact, it is this relationship that is the foundation of computer-aided design of experiments. The result has to do with the selection of the  $(r + 1)$ th point in an existing design containing  $r$  points. The coordinates to be chosen will be those such that  $|\mathbf{M}|$  is maximized. As a result, the  $(r + 1)$ th point chosen is that which produces a **conditional  $D$ -optimality**—that is, that which gives  $D$ -optimality given the initial  $r$ -point design. The result is as follows:

Given an  $r$ -point design with design and model characterized by  $\mathbf{X}_r$ , the point  $\mathbf{x}_{r+1}$  that maximizes  $|\mathbf{X}'_{r+1}\mathbf{X}_{r+1}|$  is that for which  $\mathbf{x}_{r+1}^{(m)'}(\mathbf{X}'_r\mathbf{X}_r)^{-1}\mathbf{x}_{r+1}^{(m)}$  is maximized. Hence  $\mathbf{X}_{r+1}$  is the  $\mathbf{X}$ -matrix containing the  $(r + 1)$ th point.

To understand this result, let us begin with a well-known result dealing with the determinant of a matrix. Given a square matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

then if  $\mathbf{A}_{22}$  is nonsingular,

$$|\mathbf{A}| = |\mathbf{A}_{22}||\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}|$$

and if  $\mathbf{A}_{11}$  is nonsingular,

$$|\mathbf{A}| = |\mathbf{A}_{11}||\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}|$$

Now let

$$\mathbf{X}_{r+1} = \begin{bmatrix} \mathbf{X} \\ \dots \\ \mathbf{x}'_{r+1} \end{bmatrix}$$

denote the  $\mathbf{X}$ -matrix after  $\mathbf{x}_{r+1}$  has been selected. Obviously,  $\mathbf{X}$  is  $\mathbf{X}_{r+1}$  apart from the selected point. It is easy to show that

$$(\mathbf{X}'\mathbf{X}) = \mathbf{X}'_{r+1}\mathbf{X}_{r+1} - \mathbf{x}'_{r+1}\mathbf{x}'_{r+1}$$

and hence

$$\begin{aligned} |\mathbf{X}'_{r+1}\mathbf{X}_{r+1} - \mathbf{x}'_{r+1}\mathbf{x}'_{r+1}| &= \begin{vmatrix} \mathbf{X}'_{r+1}\mathbf{X}_{r+1} & \mathbf{x}'_{r+1} \\ \mathbf{x}'_{r+1} & 1 \end{vmatrix} \\ &= |\mathbf{X}'_{r+1}\mathbf{X}_{r+1}||1 - \mathbf{x}'_{r+1}(\mathbf{X}'_{r+1}\mathbf{X}_{r+1})^{-1}\mathbf{x}'_{r+1}| \end{aligned}$$

Now the above implies that

$$|\mathbf{X}'\mathbf{X}| = |\mathbf{X}'_{r+1}\mathbf{X}_{r+1}||1 - \mathbf{x}'_{r+1}(\mathbf{X}'_{r+1}\mathbf{X}_{r+1})^{-1}\mathbf{x}_{r+1}|$$

Thus, because  $1 - \mathbf{x}'_{r+1}(\mathbf{X}'_{r+1}\mathbf{X}_{r+1})^{-1}\mathbf{x}_{r+1}$  is a scalar, we have

$$\frac{|\mathbf{X}'_{r+1}\mathbf{X}_{r+1}|}{|\mathbf{X}'\mathbf{X}|} = \frac{1}{1 - \mathbf{x}'_{r+1}(\mathbf{X}'_{r+1}\mathbf{X}_{r+1})^{-1}\mathbf{x}_{r+1}}$$

A well-known result for regression analysis relates  $\mathbf{x}'_{r+1}(\mathbf{X}'_{r+1}\mathbf{X}_{r+1})^{-1}\mathbf{x}_{r+1}$  to  $\mathbf{x}'_{r+1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_{r+1}$ . Keep in mind that the former relates to the variance of a predicted value at the location defined by  $\mathbf{x}_{r+1}$ , whereas the latter relates to the prediction variance at  $\mathbf{x}_{r+1}$  for a fitted model *before the new point was placed in the data set*. From Myers (1990) or Montgomery et al. (2012) we have

$$\mathbf{x}'_{r+1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_{r+1} = \frac{\mathbf{x}'_{r+1}(\mathbf{X}'_{r+1}\mathbf{X}_{r+1})^{-1}\mathbf{x}_{r+1}}{1 - \mathbf{x}'_{r+1}(\mathbf{X}'_{r+1}\mathbf{X}_{r+1})^{-1}\mathbf{x}_{r+1}}$$

One desires to select  $\mathbf{x}_{r+1}$  in order to maximize  $|\mathbf{X}'_{r+1}\mathbf{X}_{r+1}|$  for  $D$ -optimality. Because  $\mathbf{x}'_{r+1}(\mathbf{X}'_{r+1}\mathbf{X}_{r+1})^{-1}\mathbf{x}_{r+1}$  is a hat diagonal, we obtain  $\mathbf{x}'_{r+1}(\mathbf{X}'_{r+1}\mathbf{X}_{r+1})^{-1}\mathbf{x}_{r+1} \leq 1$ . One then must choose  $\mathbf{x}_{r+1}$  so that  $\mathbf{x}'_{r+1}(\mathbf{X}'_{r+1}\mathbf{X}_{r+1})^{-1}\mathbf{x}_{r+1}$  is maximized. From the above equation, maximizing  $\mathbf{x}'_{r+1}(\mathbf{X}'_{r+1}\mathbf{X}_{r+1})^{-1}\mathbf{x}_{r+1}$  is equivalent to maximizing  $\mathbf{x}'_{r+1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_{r+1}$ . Thus one should choose the next point  $\mathbf{x}_{r+1}$  at which the prediction variance (i.e.,  $\text{Var}[\hat{y}(\mathbf{x}_{r+1})]/\sigma^2$ ) is maximized.

The above result is not only interesting but also very useful. It essentially says that for design augmentation, the *next point* that is most appropriate in a  $D$ -optimality sense is that for which the prediction is worse when one uses the current design for prediction. In other words,  $\mathbf{x}_{r+1}$  is a type of “soft spot,” and putting this spot in the design produces the design that is conditionally  $D$ -optimal.

**Example 9.7 An Example with  $k=2$**  Consider an initial design containing a  $2^2$  factorial:

| $x_1$ | $x_2$ |
|-------|-------|
| -1    | -1    |
| -1    | 1     |
| 1     | -1    |
| 1     | 1     |

The model is a first-order model in two design variables. Now, suppose the candidate list of points are those of a  $3^2$  factorial design. The next point to be chosen is that for

which  $\mathbf{x}^{(1)'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}^{(1)}$  is largest. Here, of course,  $\mathbf{x}^{(1)'} = [1, x_1, x_2]$  and  $(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{4}\mathbf{I}_3$ , reflecting the fact that the initial design is first-order orthogonal. As a result, the expression  $\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2 = \mathbf{x}^{(1)'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}^{(1)}$  is given by

$$\begin{aligned}
 & [1, x_1, x_2] \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix} \\
 &= \frac{1}{4} [1 + x_1^2 + x_2^2]
 \end{aligned}$$

As a result, the values of  $\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  for the candidate list are

| $x_1$ | $x_2$ | $\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ |
|-------|-------|--|
| -1    | -1    | $\frac{3}{4}$                              |
| -1    | 1     | $\frac{3}{4}$                              |
| 1     | -1    | $\frac{3}{4}$                              |
| 1     | 1     | $\frac{3}{4}$                              |
| 0     | -1    | $\frac{1}{2}$                              |
| 0     | 1     | $\frac{1}{2}$                              |
| -1    | 0     | $\frac{1}{2}$                              |
| 1     | 0     | $\frac{1}{2}$                              |
| 0     | 0     | $\frac{1}{4}$                              |

Thus the best augmentation of the  $2^2$  factorial is another factorial point, and it makes no difference which one of the four is added. An interesting point here is that if the center run is added, the design remains orthogonal; but a five-run nonorthogonal design is preferable (in the  $D$ -optimal sense). This is an example that points out a shortcoming of a  $D$ -optimal design. Many experimenters would prefer to add the center run to provide some information about lack of fit. However, the optimal design will never include the center point, because it is based on the assumption is that the model is correct and  $D$ -optimality seeks to improve the estimation of terms in the model.

**Example 9.8 An Irregular Fraction with  $k = 3$**  Consider a situation in which one wishes to fit the model

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{12}x_1x_2 + \varepsilon$$

The initial design is the irregular fraction

| $x_1$ | $x_2$ | $x_3$ |
|-------|-------|-------|
| -1    | -1    | -1    |
| 1     | 1     | -1    |
| 1     | -1    | 1     |
| -1    | 1     | 1     |
| 1     | 1     | 1     |

This is a five-point design and all parameters are estimable. Suppose the list of candidate points includes the complete  $2^3$  design. Computation of  $\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  is given below for the candidate list:

| $x_1$ | $x_2$ | $x_3$ | $\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ |
|-------|-------|-------|--|
| -1    | -1    | -1    | 1.0  |
| -1    | -1    | 1     | 3.0  |
| -1    | 1     | -1    | 3.0  |
| -1    | 1     | 1     | 1.0  |
| 1     | -1    | -1    | 3.0  |
| 1     | -1    | 1     | 1.0  |
| 1     | 1     | -1    | 1.0  |
| 1     | 1     | 1     | 1.0  |

As a result, the best augmentation of the initial design involves the use of any one of the points  $(-1, -1, 1)$ ,  $(-1, 1, -1)$ , and  $(1, -1, -1)$ . Nearly all of the software algorithms make use of conditional optimization given the existing points. It is obvious how it is used for design augmentation. In this regard, one must keep in mind that a candidate list of points is used; as a result, the augmentation is chosen, sequentially one at a time, from points in the candidates list. When a design is chosen from scratch, the same result applies.

### 9.3.2 Algorithms for Computer-Generated Designs

While many computer packages now provide the capability to generate designs using one of several criteria, it is still helpful for the user to have at least a notional idea of how the algorithms approach the optimization. Two of the most popular types of algorithms are **point exchange** and **coordinate exchange**.

In the point exchange algorithm, a random design of the appropriate size is created from the set of candidate points that can estimate the assumed model. Then new designs are created by iteratively replacing individual points with the best replacement available from the candidates, conditional on the other points in the design. This is repeated for all points in the design until no improvements are possible. The entire process is repeated multiple times for different initial random designs, and the best overall design from all the starts is selected.



**Example 9.9** Consider a simple illustration of the point exchange algorithm for creating a screening design for three factors. Recall that the experimenter needs to provide information about the goals of the optimization. At the start of Section 9.3, we describe what needs to be specified:

- (i) An assumed model:  $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{12}x_1x_2 + \varepsilon$
- (ii) The required sample size: 6
- (iii) Ranges on the design variables:  $[-1, 1]$  for each factor
- (iv) A set of candidate points: the 8 2-level factorial points  $(\pm 1, \pm 1, \pm 1)$
- (v) Other constraints: no required replicates
- (vi) Criterion to optimize:  $D$ -optimality

First, we highlight a number of points from the specification of the design to be created and optimized. Note that this is a nonstandard model with just a single two-factor interaction included, perhaps chosen by the experimenter based on some additional information from experts about which terms in a standard first-order model with interactions are likely to be present. The motivation for restricting the terms in the model may have been driven by the requirement for a very small design with only six runs. In general, it is better to include additional terms in the model, in case the experts are unaware of some relationships between factors and the response. The choice of the range of design variables and the candidate points should naturally be complementary, with the candidate points filling the defined design region. Finally, since this is a screening design, selecting  $D$ -optimality that focuses on good estimation of model parameters, is a good choice for the subsequent model selection in the analysis phase.

The point exchange algorithm would begin by selecting six design points by sampling with replacement from the eight factorial points. A check would be performed to verify that the selected points can estimate the specified model; and the  $D$ -criterion,  $|M| = |X'X|/N^p$ , in Equation 9.10 is evaluated to establish a value for comparison. Perhaps for the first iteration, the original design created is as given below:

| Run | $X_1$ | $X_2$ | $X_3$ |
|-----|-------|-------|-------|
| 1   | 1     | -1    | -1    |
| 2   | 1     | 1     | -1    |
| 3   | -1    | -1    | -1    |
| 4   | 1     | -1    | 1     |
| 5   | -1    | 1     | -1    |
| 6   | 1     | -1    | -1    |

Then for the first row in the design matrix, a comparison between the designs with the current candidate point  $(1, -1, -1)$  and each of the seven alternate candidate points is made, with  $|M|$  evaluated each time. The candidate that improves the criterion the most is selected and becomes the new first run. Next, the second run  $(1, 1, -1)$  is evaluated the same way, and potentially updated with a better alternative. This process is repeated for all of the six runs of the design, and then the algorithm returns to the first run, and so on. The search continues until no improvements can be made for this design.

This process is then repeated for multiple random starting designs, with the best design from each start being compared to the best  $|M|$  value found so far. The final “optimal” design is the best found from all of the iterations of the algorithm. When JMP was used to generate a design satisfying the specified requirements, the resulting design is listed below:

| Run | $X_1$ | $X_2$ | $X_3$ |
|-----|-------|-------|-------|
| 1   | -1    | -1    | 1     |
| 2   | -1    | 1     | 1     |
| 3   | -1    | -1    | -1    |
| 4   | 1     | -1    | 1     |
| 5   | 1     | -1    | -1    |
| 6   | 1     | 1     | -1    |

It is helpful to note that the “optimal design” is not unique. If the roles of  $-1$  and  $+1$  for one or more columns were reversed, the resulting design would still have the same  $D$ -efficiency properties. In addition, there may be multiple designs that all have the same best  $D$ -efficiency, and so it can be helpful to consider supplementary characteristics to distinguish between alternatives. See Robinson and Anderson-Cook (2011) for more discussion about selecting between tied  $D$ -optimal designs.

The coordinate exchange algorithm is a candidate-free approach, which begins similarly to the point exchange algorithm with a random design of the appropriate size. Then instead of focusing on a row of the design matrix, it examines each of the entries individually. Each coordinate for each factor for each design point is sequentially adjusted to its optimal value within the range of that factor, conditional on all other coordinates and points in the design. If we assume the same initial design as in Example 9.9, this would mean that the algorithm would look at the entry for  $X_1$  in run 1, and try to improve the design while keeping all of the other entries fixed. Then it would move to the entry for  $X_2$  in run 1, and so on, throughout all 18 entries in the design matrix. This is repeated for all coordinates in the design until no further improvements exist. As with the point exchange algorithm, multiple random initial designs are considered and the best overall design from all starts is chosen.

The algorithms for  $D$ - or  $I$ -optimality leverage computationally efficient methods for updating the effect of changing a single point within a design. Hence, these algorithms can be used quickly and easily, and can consider a large number of initial designs in a timely manner. Until recently,  $G$ -optimality was considerably more time-consuming and computationally intensive to use as the criterion for optimal designs. Rodriguez et al. (2010) utilize Brent’s algorithm [Brent (1973)] to dramatically improve the speed with which these designs can be generated.

Because construction of an alphabetically optimal design is essentially a combinatorial optimization problem, methods for solving these types of problems could also be employed. Indeed, some authors have used branch and bound techniques [Welch (1982)] to construct optimal designs. Simulated annealing [Meyers and Nachtsheim (1988) and Drain et al. (2004)] and genetic algorithms [Heredia-Langner et al. (2003) and Lin et al. (2015)] have also been used.

It should be strongly emphasized that the goal of the computer-generated design algorithms is to produce a design that gives approximately the best performance for the criterion selected. Different packages may, indeed, produce different designs with the same input information. Additionally, there are often choices about the running of the algorithm in the software, such as the number of random starts to use, which should be examined. For most experiments, the cost, time, and effort to set up and run the experiment is considerably greater than spending some additional time with the computer software to ensure that the best possible design has been found to match the study goals.

### 9.3.3 Comparison of $D$ -, $G$ -, and $I$ -Optimal Designs

Since efficient algorithms exist for finding optimal designs for each of the three primary criteria for good estimation ( $D$ ) and prediction ( $G$  and  $I$ ), it is helpful to look at a comparison of the performance of the best designs relative to each other for all three criteria. Rodriguez et al. (2010) examine designs for second-order models in a cuboidal design space. Table 9.3 shows the relative efficiency of each design, where clearly the  $D$ -optimal design will be 100%  $D$ -efficient, as will each of the other optimal designs for their own criterion. What is of interest is to see how the designs perform for the other two criteria, as this will help show their robustness if multiple characteristics are being considered when selecting which designed experiment to perform.

We note a number of patterns in the table. First, all of the designs are more similar for their  $D$ -efficiency than for the other criteria. In other words,  $G$ - and  $I$ -optimal designs tend to have  $D$ -efficiencies that are higher than when we consider the performance of all three designs for  $G$ - or  $I$ -efficiency. This says that if we focus exclusively on  $G$ - or  $I$ -optimality while generating the design, we will likely still see pretty good performance on  $D$ -efficiency.

If the focus is on characteristics of the prediction variance, we see that there are substantial trade-offs between  $G$ - and  $I$ -efficiency. This should make some intuitive sense, as whether we focus on an extreme of a distribution (like the maximum, for  $G$ -efficiency) or a measure of its center (like the average, for  $I$ -efficiency) reflects very different emphases on what to optimize. Typically, the center of distributions has been a more common focus in the study of statistics, so this would suggest that  $I$ -optimality might be a better choice. However, protecting against the worst-case prediction variance and keeping that more moderate, also has some appeal, since often the user does not know a priori where there will be interest in making new predictions.

There is no consistent pattern of which of the  $G$ - or  $I$ -optimal designs performs better on the other criterion, and looking at the single number summaries is not that helpful for understanding what the tradeoffs between the designs look like. A new plot, the Variance Ratio Fraction of Design Space (VRFDS) plot, looks at the values of the prediction variance at a number of locations for different designs. For locations throughout the design space, the prediction variance values are calculated for all of the designs to be compared, and the ratios of these values at each location are calculated, with one design being chosen as the reference design. Hence, for our three designs, we might select the  $G$ -optimal design as the reference distribution and then calculate

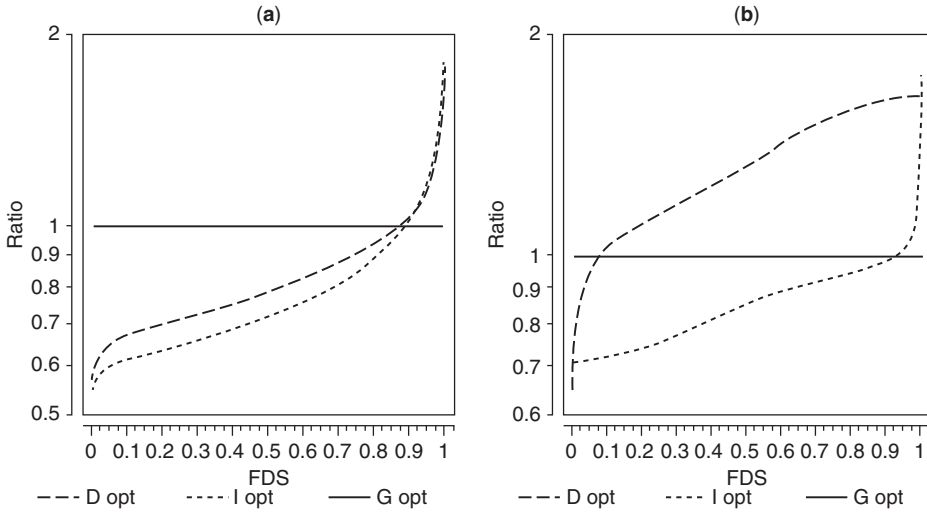
$$VR_D = \frac{PV_D(\mathbf{x})}{PV_G(\mathbf{x})} \quad \text{and} \quad VR_I = \frac{PV_I(\mathbf{x})}{PV_G(\mathbf{x})}$$

**TABLE 9.3 Comparison of Optimal Designs for Second-Order Models in a Cuboidal Region**

| <i>n</i>  | Optimal Design | <i>G</i> -eff. | <i>D</i> -eff. | <i>I</i> -eff. | <i>n</i>  | Optimal Design | <i>G</i> -eff. | <i>D</i> -eff. | <i>I</i> -eff. |
|-----------|----------------|----------------|----------------|----------------|-----------|----------------|----------------|----------------|----------------|
| 2 Factors |                |                |                |                | 3 Factors |                |                |                |                |
| 6         | G              | 100.00         | 94.60          | 92.40          | 10        | G              | 100.00         | 91.40          | 79.06          |
|           | D              | 71.82          | 100.00         | 93.67          |           | D              | 77.40          | 100.00         | 96.84          |
|           | I              | 57.96          | 91.17          | 100.00         |           | I              | 69.38          | 97.56          | 100.00         |
| 7         | G              | 100.00         | 88.06          | 75.10          | 11        | G              | 100.00         | 93.27          | 74.97          |
|           | D              | 78.29          | 100.00         | 86.93          |           | D              | 78.30          | 100.00         | 70.97          |
|           | I              | 56.41          | 85.11          | 100.00         |           | I              | 63.67          | 91.51          | 100.00         |
| 8         | G              | 100.00         | 97.55          | 81.05          | 12        | G              | 100.00         | 93.04          | 69.39          |
|           | D              | 75.61          | 100.00         | 88.45          |           | D              | 69.33          | 100.00         | 70.38          |
|           | I              | 64.94          | 89.70          | 100.00         |           | I              | 59.91          | 87.66          | 100.00         |
| 9         | G              | 100.00         | 93.86          | 81.28          | 13        | G              | 100.00         | 98.50          | 68.90          |
|           | D              | 95.87          | 100.00         | 94.76          |           | D              | 95.78          | 100.00         | 72.77          |
|           | I              | 70.44          | 90.71          | 100.00         |           | I              | 65.17          | 86.80          | 100.00         |
| 10        | G              | 100.00         | 94.31          | 75.45          | 14        | G              | 100.0          | 99.95          | 98.21          |
|           | D              | 88.40          | 100.00         | 84.29          |           | D              | 93.28          | 100.00         | 76.78          |
|           | I              | 87.54          | 97.37          | 100.00         |           | I              | 53.08          | 86.10          | 100.00         |
| 11        | G              | 100.00         | 93.93          | 70.88          | 15        | G              | 100.00         | 99.57          | 90.17          |
|           | D              | 82.25          | 100.00         | 78.14          |           | D              | 93.16          | 100.00         | 72.63          |
|           | I              | 79.77          | 92.81          | 100.00         |           | I              | 99.65          | 97.14          | 100.00         |
| 12        | G              | 100.00         | 89.21          | 73.90          | 16        | G              | 100.00         | 96.94          | 95.39          |
|           | D              | 78.30          | 100.00         | 75.78          |           | D              | 92.14          | 100.00         | 71.03          |
|           | I              | 74.39          | 87.56          | 100.00         |           | I              | 98.97          | 93.65          | 100.00         |
| 4 Factors |                |                |                |                | 5 Factors |                |                |                |                |
| 15        | G              | 100.00         | 98.01          | 75.89          | 21        | G              | 100.00         | 98.89          | 65.20          |
|           | D              | 98.28          | 100.00         | 78.23          |           | D              | 92.86          | 100.00         | 63.04          |
|           | I              | 76.44          | 87.01          | 100.00         |           | I              | 54.35          | 78.26          | 100.00         |
| 20        | G              | 100.00         | 93.75          | 73.40          | 26        | G              | 93.08          | 99.34          | 80.86          |
|           | D              | 95.18          | 100.00         | 65.63          |           | D              | 86.41          | 100.0          | 61.06          |
|           | I              | 82.11          | 85.48          | 100.00         |           | I              | 64.50          | 82.69          | 100.00         |
| 24        | G              | 99.97          | 96.59          | 87.00          | 30        | G              | 100.00         | 99.93          | 60.04          |
|           | D              | 90.71          | 100.00         | 61.82          |           | D              | 98.65          | 100.00         | 60.15          |
|           | I              | 55.38          | 87.79          | 100.00         |           | I              | 71.06          | 87.08          | 100.00         |

at a large number of design locations. The FDS plot is then obtained by sorting the ratios from smallest to largest. Recall that the goal is have small prediction variance values throughout the design space, so the reference design is better if the ratio is greater than one, and the comparison design (here *D*- or *I*-optimal designs) if the ratio is less than one.

For example, consider two scenarios, say three factors with 10 and 15 runs, respectively. For the three factor and 10-run design, the *G*-optimal design seems more robust across



**Figure 9.14** Variance Ratio Fraction of Design Space plots for three-factor designs fitting a second-order model for (a) 10 and (b) 15 runs.

criteria with an *I*-efficiency of 79.06%, while the *I*-optimal design with *G*-efficiency of 69.38%. The reverse pattern is true for the 3 factor and 15-run design scenario has the *G*-optimal design with *I*-efficiency of 90.17%, and the *I*-optimal design with *G*-efficiency of 99.65%. Figure 9.14 shows the VRFDS plots for the two scenarios, with all three designs shown and the *G*-optimal design selected as the reference design.

From Figure 9.14, it is clear that the single number summary of the design may be too simplistic. For both scenarios, for the majority of the design space (nearly 90% or 95% of the region), the *I*-optimal design has a smaller prediction variance than the *G*-optimal design. Only for a very small fraction of the design space does the *G*-optimal design perform better. This is striking, because our impression of the three-factor 10-run design from Table 9.3 was that it seemed that the *G*-optimal design was more robust across the criteria. The performance of the *D*-optimal design differs much more between the scenarios, but perhaps it is to be expected that it would be less consistent, given that it was chosen as optimal with a focus on estimation of the model parameters.

As noted by Rodriguez et al. (2010), this pattern of better prediction variance from the *I*-optimal designs for the vast majority of design locations is consistent across the different scenarios included in Table 9.3. Hence, justifying the choice of *G*-optimal designs as a preferred option is difficult and in general, if good prediction in the design region is a primary goal of the experiment, then *I*-optimal designs generally have better overall performance. Since the *D*-optimal designs have quite inconsistent results relative to the *I*-optimal designs, it is best to think carefully about whether estimation of model parameters or prediction are more important for a particular experiment.

### 9.3.4 Illustrations Involving Computer-Generated Design

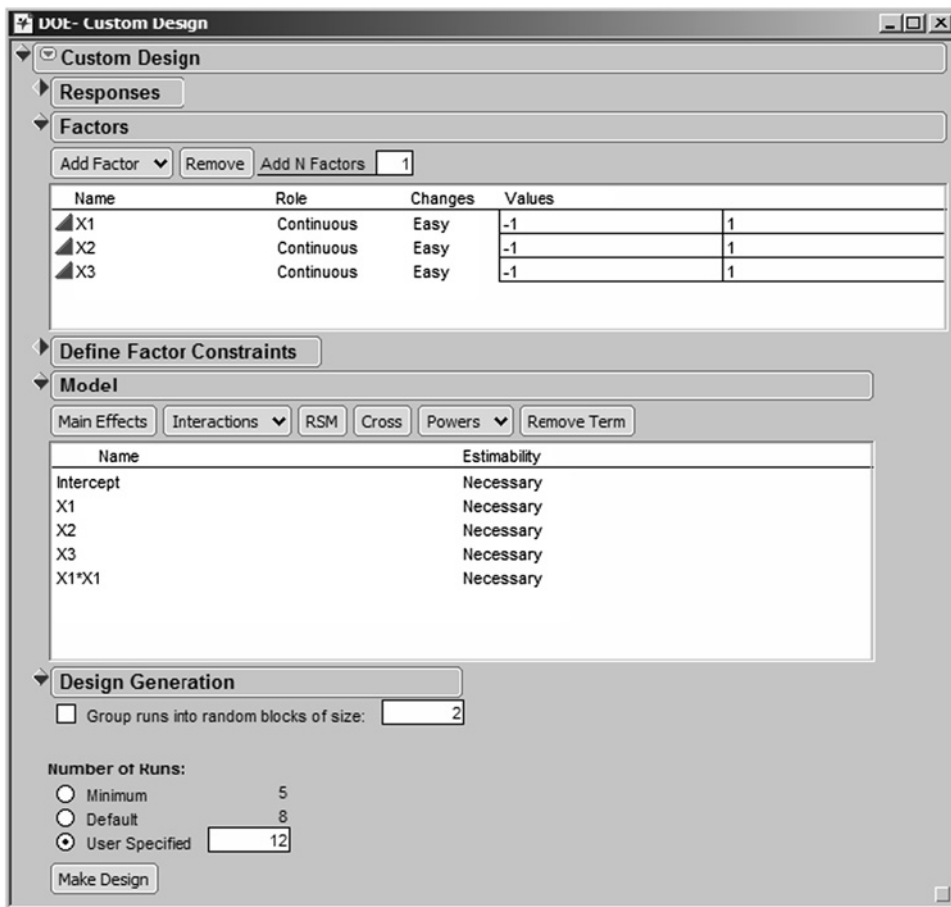
In this section we focus on a few illustrations of computer-generated designs with emphasis on the use of JMP and Design-Expert. This does not mean to imply that these are the only or even the best computer packages. Our first example is a fairly simple one where, based on

our discussions in this and other chapters, the reader likely could have correctly identified the most appropriate design.

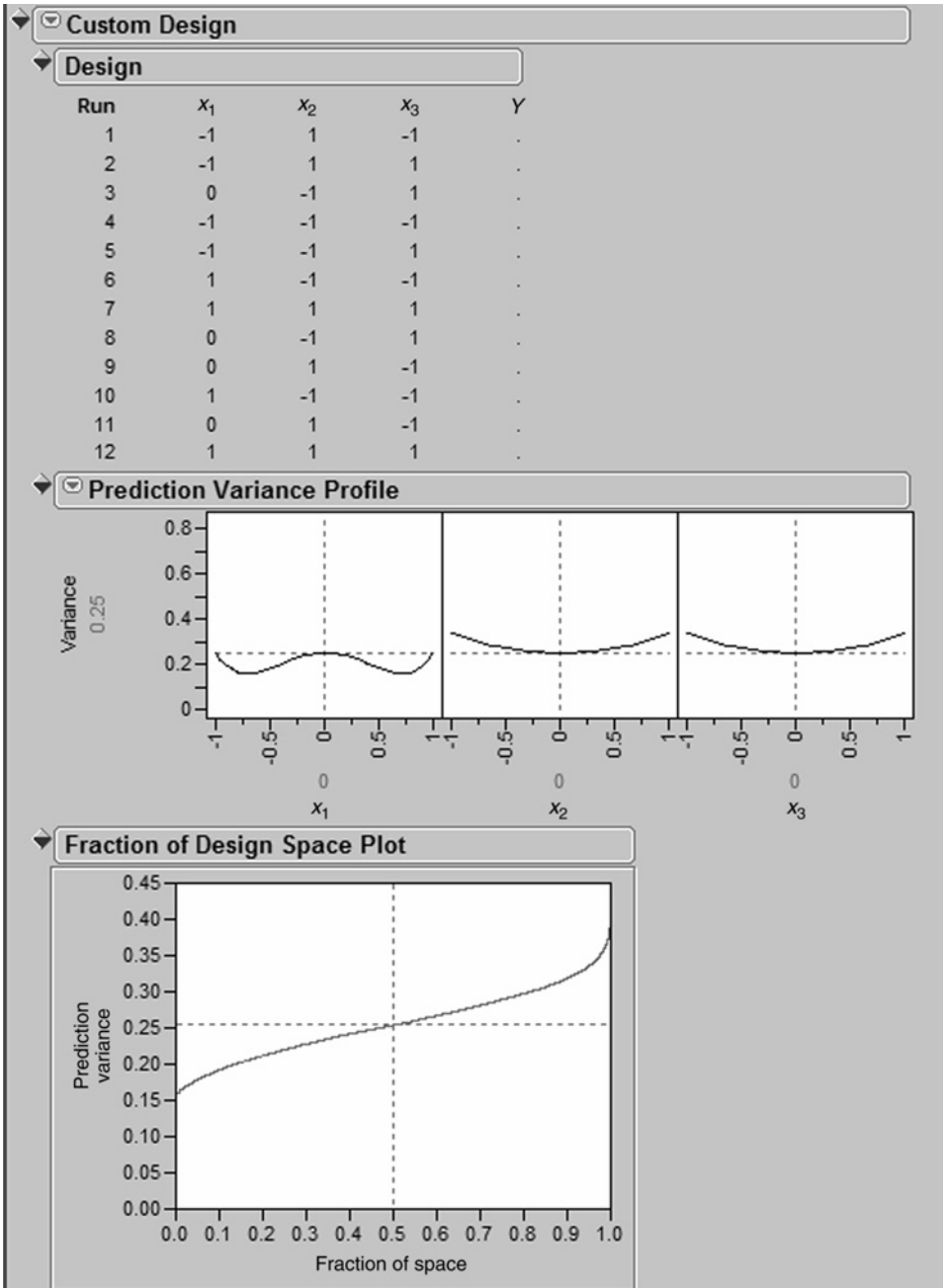
**Example 9.10 *D*- and *I*-Optimal Designs for a Nonstandard Model** Suppose there are  $k = 3$  design variables, and an intelligent “guess” gives the appropriate model as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{11} x_1^2 + \varepsilon$$

The ranges being considered by the researcher are  $[-1, +1]$  (coded) for all three variables and the experimenter wishes to have 12 runs. Using the “Custom Design” option under “DOE” in JMP, we specify the model with three continuous factors and the model given above. Figure 9.15 shows the dialog boxes for creating the design, while Fig. 9.16 shows the created design. Under the red arrow for Custom Design show in Fig. 9.15, the “Optimality Criterion” selected was “Make *D*-Optimal Design” and the “Number of starts” was set to 100.



**Figure 9.15** Custom Design dialog box from JMP to create design with 12 runs for the nonstandard model.



**Figure 9.16** Result of Custom Design for *D*-optimal design for the nonstandard model, with table of design points, Prediction Variance Profiler and FDS plot.

Note that the design in the output is a  $3 \times 2 \times 2$  factorial. As the reader might expect, three levels were selected for  $x_1$ , since the experimenter wishes to estimate a quadratic term, while only two levels were selected for each of  $x_2$  and  $x_3$ . The design chosen is  $D$ -optimal. It turns out that the  $G$ -efficiency of the design is 100%. In fact, the maximum value of  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  is five, the total number of model parameters. This can be evaluated by using the FDS plot or the Prediction Variance Profiler to find that the maximum value of  $v(\mathbf{x}) = \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  is 0.4167. Note that JMP gives the unscaled prediction variance in the output which is not scaled by the sample size, and  $5 = 12 \times 0.4167$ .

As a result the design is  $G$ -optimal. Note that the  $I$ -efficiency of the design is difficult to assess, since this value is not provided in the output. By looking at the 0.5 fraction of the FDS plot, we can estimate the median value of the prediction variance, but this value is likely not a good approximation of the average prediction variance since the distribution of the prediction variance throughout the design space is skewed.

If instead of focusing on  $D$ -optimality, we had been interested in using the  $I$ -optimality criterion, we could specify this option in JMP with the “Optimality Criterion” selected as “Make  $I$ -Optimal Design.” The design obtained is as follows:

| $x_1$ | $x_2$ | $x_3$ |
|-------|-------|-------|
| -1    | -1    | 1     |
| -1    | 1     | 1     |
| -1    | 1     | -1    |
| 0     | 1     | 1     |
| 0     | 1     | 1     |
| 0     | -1    | -1    |
| 0     | -1    | -1    |
| 0     | 1     | -1    |
| 0     | 1     | -1    |
| 1     | 1     | 1     |
| 1     | 1     | -1    |
| 1     | -1    | 1     |

This design places more emphasis on good prediction at the center of the design space with additional points allocated for  $x_1 = 0$ . This comes at the expense of the worst prediction in the region being inflated to 0.6667, which would lead to a  $G$ -efficiency of  $5/(12 \times 0.6667) = 62.5\%$ . This result matches what was noted in the previous section that optimizing good average prediction in the entire region often comes at the cost of making the worst prediction variance larger. In addition, using an  $I$ -optimal design improves prediction for large portions of the design region.

**Example 9.11 Use of Additional Runs** Consider the same scenario as in the previous example except that the user can afford to use  $N = 15$  runs. Using the “Augment Design” option under “DOE” in JMP, we can add three additional runs to the design created in Example 9.10. Figure 9.17 shows the dialog box used to create the new design.

Figure 9.18 shows the resulting design, with the three additional runs. Note that we have adjusted the Prediction Variance Profiler to show the maximum prediction variance value at one of the corners of the design space. This design has a lower maximum  $v(\mathbf{x})$  value



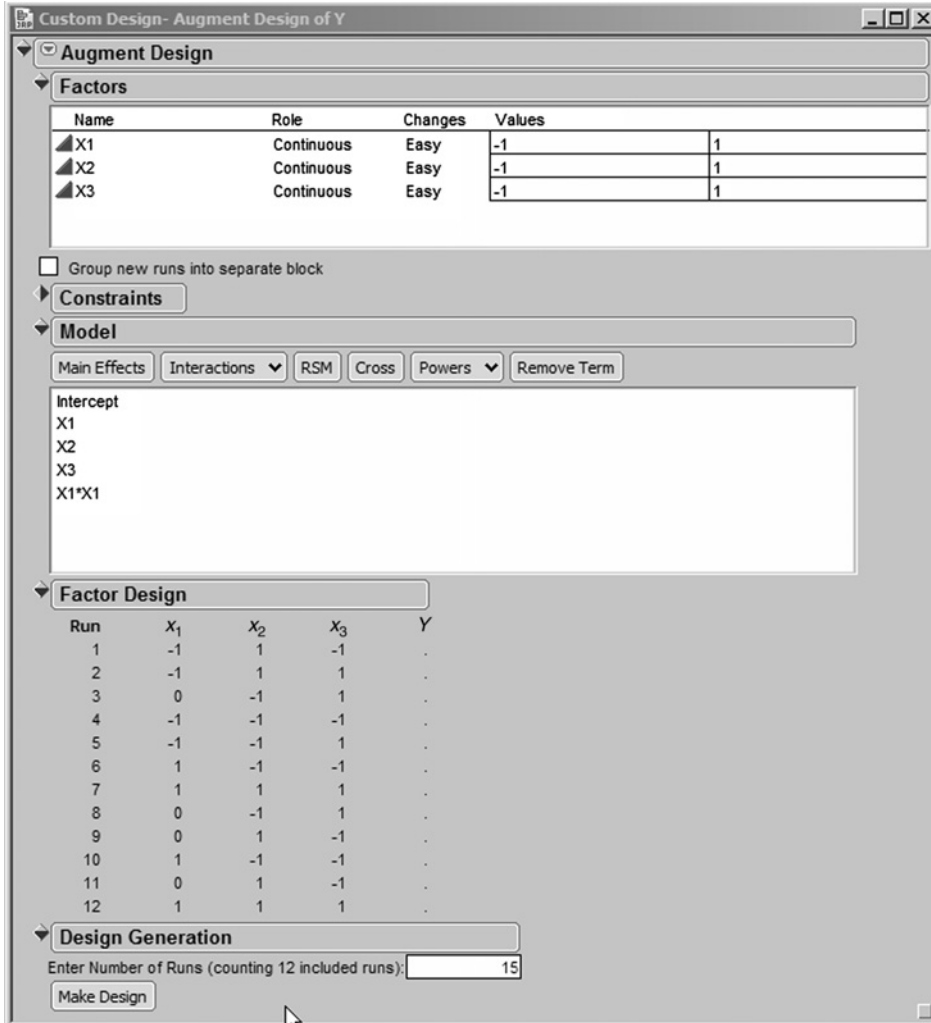


Figure 9.17 Dialog box to augment the 12-run design of Example 9.10 to a 15-run design.

of 0.375, which makes this a better design for prediction. However, if we are using scaled prediction variance, then the maximum  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2 = 15 \times 0.375 = 5.625$ . This gives a  $G$ -efficiency of  $5/5.625 = 89\%$ . Hence this design is not superior to the 12-run design on a per observation basis, but will improve prediction if we consider the unscaled prediction variance.

The use of  $D$ -efficiency under constraints on lack-of-fit and pure error degrees of freedom via computer packages is a continuing attempt to make the “automated design” concept practical. Quite often the constraints revolve around total sample size, levels, and so on. However, there are algorithms that generate a most  $D$ -efficient design where the constraints are built around pure error and lack-of-fit degrees of freedom. For example, suppose that in a  $k = 3$  situation it is of interest to design an experiment with two lack-of-fit degrees of freedom and two pure error degrees of freedom. For a second-order model this corresponds

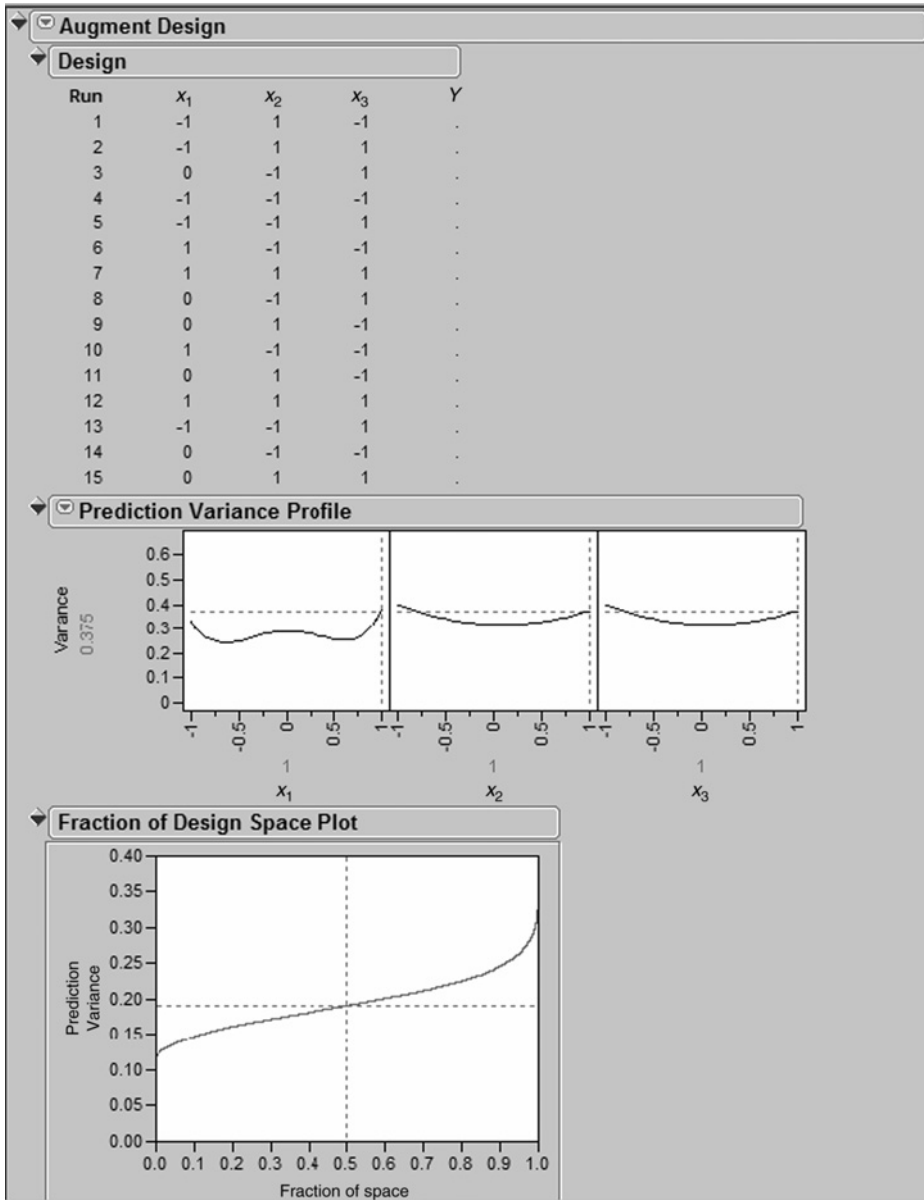


Figure 9.18 Augmented design for nonstandard model with  $N = 15$ .

to 12 distinct design points and 15 total runs. The resulting design, produced by Design-Expert, is given in Table 9.4.

Note from Table 9.4 that the design allows a single center run and has replication at two distinct factorial points. While we are accustomed to seeing designs with center runs replicated, there are certainly advantages to having replication at *different points* so some examination of the homogeneous variance assumption can be made.

**TABLE 9.4 Output from Design-Expert for  $D$ -Optimal Design with  $k = 3$ , Two Lack-of-Fit Degrees of Freedom, and Two Pure Error Degrees of Freedom**

| $x_1$ | $x_2$ | $x_3$ |
|-------|-------|-------|
| -1.0  | -1.0  | -1.0  |
| 1.0   | -1.0  | -1.0  |
| 1.0   | -1.0  | -1.0  |
| 0.0   | 0.0   | -1.0  |
| -1.0  | 1.0   | -1.0  |
| -1.0  | 1.0   | -1.0  |
| 1.0   | 1.0   | -1.0  |
| -1.0  | 0.0   | 0.0   |
| 0.0   | 0.0   | 0.0   |
| 0.0   | 1.0   | 0.0   |
| -1.0  | -1.0  | 1.0   |
| -1.0  | -1.0  | 1.0   |
| -1.0  | 1.0   | 1.0   |
| 1.0   | 1.0   | 1.0   |
| -0.5  | -0.5  | -0.5  |

**Example 9.12 A Nonstandard Model with  $k = 4$  Variables** In this example we provide another illustration of the use of computer-generated designs. This example serves as an illustration of some of the *disadvantages* of computer-generated design and design optimality in general.

Suppose we have four design variables and the input involves three levels on each, coded to  $-1, 0, +1$ . Thus, the candidate list will be the  $3^4 = 81$  design points. Suppose the practitioner has difficulty choosing a model for purpose of design construction but decides on

$$E(y) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \beta_{12}x_1x_2 + \beta_{23}x_2x_3 + \beta_{11}x_1^2 + \beta_{44}x_4^2 \tag{9.24}$$

The practitioner can afford  $N = 25$  runs. The  $D$ -optimal design is

| Observation | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|-------------|-------|-------|-------|-------|
| 1           | -1    | -1    | -1    | -1    |
| 2           | -1    | -1    | -1    | 0     |
| 3           | -1    | -1    | -1    | 0     |
| 4           | -1    | -1    | 1     | -1    |
| 5           | -1    | -1    | 1     | 1     |
| 6           | -1    | 1     | -1    | -1    |
| 7           | -1    | 1     | -1    | 0     |
| 8           | -1    | 1     | 1     | 1     |
| 9           | -1    | 1     | 1     | 1     |
| 10          | 0     | -1    | -1    | 1     |

(continued)

| Observation | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|-------------|-------|-------|-------|-------|
| 11          | 0     | -1    | 1     | 0     |
| 12          | 0     | -1    | 1     | 1     |
| 13          | 0     | 1     | -1    | -1    |
| 14          | 0     | 1     | -1    | 0     |
| 15          | 0     | 1     | 1     | -1    |
| 16          | 0     | 1     | 1     | 0     |
| 17          | 1     | -1    | -1    | -1    |
| 18          | 1     | -1    | -1    | 0     |
| 19          | 1     | -1    | -1    | 1     |
| 20          | 1     | -1    | 1     | -1    |
| 21          | 1     | -1    | 1     | 0     |
| 22          | 1     | 1     | -1    | 1     |
| 23          | 1     | 1     | -1    | 1     |
| 24          | 1     | 1     | 1     | -1    |
| 25          | 1     | 1     | 1     | 0     |

The  $D$ -efficiency is 62.35%, and the  $G$ -efficiency is 94.13%.

It is of interest to compare this design with a standard design. In fact, a CCD with  $\alpha = 1.0$  and one center run satisfies the  $N = 25$  requirements. Of course, the CCD is totally symmetric, requires three levels of each variable, and estimates coefficients for all six two-factor interactions equally well. In addition, all pure quadratic coefficients are estimated equally well. As one would expect, the CCD is not as efficient for the incomplete second-order model listed above. In fact, the efficiencies are

$$D_{\text{eff}} = 52.13\%$$

$$G_{\text{eff}} = 94.09\%$$

How often, however, is the practitioner certain of the model prior to designing an experiment, particularly in an RSM scenario when most (or all) factors are continuous? Suppose we consider a situation in which, after the data are collected, the model that is fitted to the data is given by

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{34} x_3 x_4 \quad (9.25)$$

Of course, the constructed design is no longer  $D$ -optimal. In fact, a highly efficient two-level design could be constructed with the computer. However, one must live with the design chosen. Now, the CCD was also not designed for the model of Equation 9.25, particularly because there are no quadratic terms. The efficiencies are:

|                    | $D$   | $A$   | $G$   |
|--------------------|-------|-------|-------|
| Computer-generated | 75.88 | 67.86 | 82.07 |
| CCD                | 72.96 | 72.31 | 85.04 |

As a result, the CCD is just as efficient as (if not more so than) the computer-generated design. Now, suppose the actual model, unknown to the practitioner, involves a quadratic

**TABLE 9.5** A *D*-Optimal Design for the Constrained Region in Fig. 9.1

| Standard Order | $x_1$ | $x_2$ | $y$   |
|----------------|-------|-------|-------|
| 1              | -0.50 | -1.00 | 173.9 |
| 2              | 1.00  | 0.00  | 195.9 |
| 3              | -0.08 | -0.08 | 201.2 |
| 4              | -1.00 | 1.00  | 199.3 |
| 5              | 1.00  | -1.00 | 190.3 |
| 6              | 0.00  | 1.00  | 198.5 |
| 7              | -1.00 | 0.25  | 190.5 |
| 8              | 0.25  | -1.00 | 179.5 |
| 9              | -1.00 | -0.50 | 176.7 |
| 10             | 1.00  | 0.00  | 198.1 |
| 11             | 0.00  | 1.00  | 200.5 |
| 12             | -0.08 | -0.08 | 197.1 |

term in either  $x_2$  or  $x_3$ . Here, the computer-generated design of Equation 9.24 is a **singular design**; that is, the coefficients of  $x_2^2$  and  $x_3^2$  are not estimable. The point here is that the CCD has a high *D*-efficiency and *G*-efficiency for the full second-order model and is often reasonably robust to model mis-specification, in contrast with a computer-generated design. One cannot always be certain that a computer-generated *D*-optimal design will allow estimation of all coefficients if the model that is eventually fitted differs from that used in the input information.

**Example 9.13 The Adhesive Bonding Experiment** Consider the adhesive bonding experiment described in Section 9.2 that led to the constrained region of interest in Fig. 9.1. Recall that the two variables are  $x_1$  = amount of adhesive and  $x_2$  = cure temperature, and that the feasible design region is

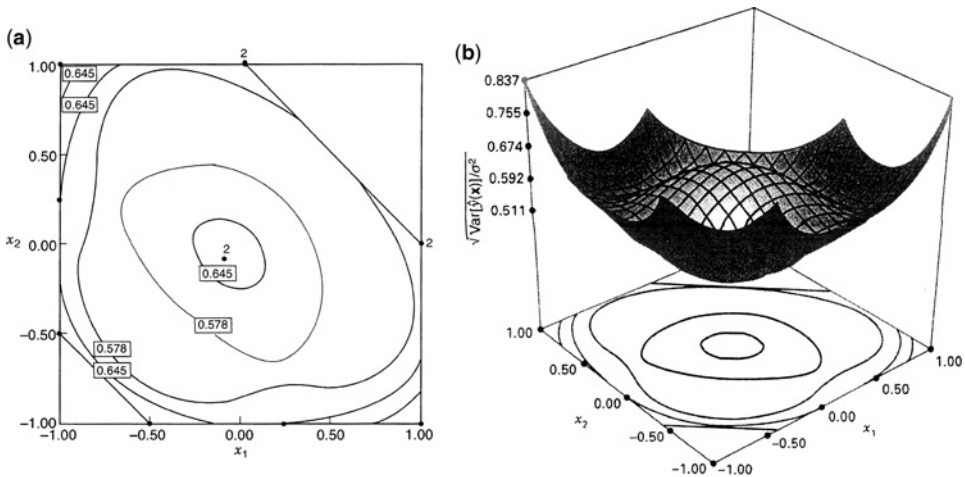
$$-1.5 \leq x_1 + x_2 \leq 1.0$$

Suppose that the response of interest is the pulloff force and we want to fit a second-order model to this response. Table 9.5 and Fig. 9.19a show a 12-run *D*-optimal design for this problem, generated with Design-Expert. We required that  $N=12$ , and that there would be three replicated runs to provide an estimate of pure error. Notice that one of the replicates is at the center of the region, and two are at corners, or **vertices**. Figure 9.19a shows the contours of  $\sqrt{\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2}$  for this design, and Fig. 9.19b shows the corresponding response surface.

Table 9.6 is a summary of the fit of a second-order model for the design in Table 9.5 using Design-Expert. The fitted model is

$$\hat{y} = 199.10 + 4.51x_1 + 9.56x_2 - 9.57x_1x_2 - 6.19x_1^2 - 9.00x_2^2$$

Figure 9.20 shows the contour plot and the response surface for the estimated response for this model. The point giving the maximum pulloff force is  $x_1 = 0.08$  and  $x_2 = 0.52$ .



**Figure 9.19** A  $D$ -optimal design for the constrained design region in Fig. 9.1. (a) The design and contours of constant  $\sqrt{\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2}$ . (b) The response surface plot.

### 9.3.5 Computer-Generated Designs Involving Qualitative Variables

Quite often in practice, response surface studies need to be conducted even though at least one factor is qualitative in nature. For example, in a manufacturing environment, there may be different machines or lines that nominally do the same function, or there may be different suppliers of a given ingredient. In these cases there are set levels of the factor, and the factor is categorical. Recall in Chapter 3 that qualitative factors required changes to the design, in particular for center runs where center runs were included at each level of the qualitative factors, as in Fig. 3.24.

In an RSM study involving qualitative variables, it stands to reason that the goal of the experiment for these variables is no different from that of the quantitative variables. That is, they must be included in the model and designs, and they must be involved in any response prediction or optimization using the model. The reader should bear in mind that *qualitative variables in an RSM study are not like blocks*. They are not nuisance factors. Indeed one must allow for possible interaction between qualitative factors and standard RSM terms involving quantitative factors.

Careful consideration should be given to what relationships between the other factors and the response are anticipated to remain the same across the levels of the qualitative factor, and which are expected to change. If no interaction terms are included between the qualitative and quantitative factors, then the assumption is that the relationship between quantitative factors and the response remains the same, but with just a fixed shift between the levels of the qualitative factor. With the addition of interaction terms, the model can accommodate increasing departures from this assumption. If all interaction terms between the qualitative and quantitative factors are included, then completely different relationships between the quantitative factors and the response can be modeled for each level of the qualitative factor.

Consider a simple example with two quantitative factors ( $x_1, x_2$ ) and one qualitative factor with two levels. One possible model for the response might be

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \delta_3 z_3 + \varepsilon$$

**TABLE 9.6 The Second-Order Model for the Pulloff Force**

Response: Force  
 \*\*\*WARNING: The Cubic Model is Aliased!\*\*\*

| Source   | Sum of Squares | DF        | Sequential Model Sum of Squares |                     |                    | F-Value   | Prob > F  |
|--|----------------|-----------|---------------------------------|---------------------|--------------------|-----------|-----------|
|  |                |           | Mean Square                     | R-Squared           | Adjusted R-Squared |           |           |
| Mean   | 4.414E+005     | 1         | 4.414E+005                      |                     |                    |           |           |
| Linear   | 809.80         | 2         | 404.90                          |                     | 14.96              | 0.0014    |           |
| 2FI  | 30.83          | 1         | 30.83                           |                     | 1.16               | 0.3130    |           |
| Quadratic  | 165.87         | 2         | 82.94                           |                     | 10.61              | 0.0107    | Suggested |
| Cubic  | 34.10          | 3         | 11.37                           |                     | 2.66               | 0.2216    | Aliased   |
| Residual   | 12.83          | 3         | 4.28                            |                     |                    |           |           |
| Total  | 4.425E+005     | 12        | 36871.83                        |                     |                    |           |           |
| <i>Lack of Fit Tests</i>                                       |                |           |                                 |                     |                    |           |           |
| Linear   | 230.80         | 6         | 38.47                           |                     | 9.06               | 0.0496    |           |
| 2FI  | 199.97         | 5         | 39.99                           |                     | 9.36               | 0.0475    |           |
| Quadratic  | 34.10          | 3         | 11.37                           |                     | 2.66               | 0.2216    | Suggested |
| Cubic  | 0.000          | 0         |                                 |                     |                    |           | Aliased   |
| Pure Error   | 12.82          | 3         | 4.27                            |                     |                    |           |           |
| <i>Model Summary Statistics</i>                                |                |           |                                 |                     |                    |           |           |
| Source   | Std. Dev.      | R-Squared | Adjusted R-Squared              | Predicted R-Squared | PRESS              |           |           |
| Linear   | 5.20           | 0.7687    | 0.7173                          | 0.6494              | 369.32             |           |           |
| 2FI  | 5.16           | 0.7980    | 0.7222                          | 0.6982              | 317.96             |           |           |
| Quadratic  | 2.80           | 0.9555    | 0.9183                          | 0.8260              | 183.32             | Suggested |           |
| Cubic  | 2.07           | 0.9878    | 0.9554                          |                     | +                  | Aliased   |           |
| +Cases(s) with leverage of 1.0000: PRESS statistic not defined |                |           |                                 |                     |                    |           |           |

(continued)

TABLE 9.6 (Continued)

| Response Force |                | ANOVA for Response Surface Quadratic Model<br>Analysis of variance table [Partial sum of squares] |                |         |          |  |
|----------------|----------------|---|----------------|---------|----------|--|
| Source         | Sum of Squares | DF  | Mean Square    | F-Value | Prob > F |  |
| Model          | 1006.51        | 5   | 201.30         | 25.74   | 0.0006   |  |
| A              | 110.69         | 1   | 110.69         | 14.15   | 0.0094   |  |
| B              | 496.83         | 1   | 496.83         | 63.53   | 0.0002   |  |
| A <sup>2</sup> | 72.86          | 1   | 72.86          | 9.32    | 0.0224   |  |
| B <sup>2</sup> | 154.39         | 1   | 154.39         | 19.74   | 0.0044   |  |
| AB             | 137.55         | 1   | 137.55         | 17.59   | 0.0057   |  |
| Residual       | 46.92          | 6   | 7.82           |         |          |  |
| Lack of Fit    | 34.10          | 3   | 11.37          | 2.66    | 0.2216   |  |
| Pure Error     | 12.82          | 3   | 4.27           |         |          |  |
| Cor Total      | 1053.43        | 11  |                |         |          |  |
| Std. Dev.      | 2.80           |   |                |         |          |  |
| Mean           | 191.79         |   |                |         |          |  |
| C.V.           | 1.46           |   |                |         |          |  |
| PRESS          | 183.32         |   |                |         |          |  |
|                |                |   | R-Squared      | 0.9555  |          |  |
|                |                |   | Adj R-Squared  | 0.9183  |          |  |
|                |                |   | Pred R-Squared | 0.8260  |          |  |
|                |                |   | Adeq Precision | 14.01   |          |  |



95% CI

| Factor         | Coefficient Estimate | DF | Standard Error | 95% CI |        |  | VIF  |
|----------------|----------------------|----|----------------|--------|--------|--|------|
|                |                      |    |                | Low    | High   |  |      |
| Intercept      | 199.10               | 1  | 1.84           | 194.61 | 203.60 |  |      |
| A-Amount       | 4.51                 | 1  | 1.20           | 1.58   | 7.45   |  | 1.16 |
| B-Temp         | 9.56                 | 1  | 1.20           | 6.62   | 12.49  |  | 1.16 |
| A <sup>2</sup> | -6.19                | 1  | 2.03           | -11.14 | -1.23  |  | 1.43 |
| B <sup>2</sup> | -9.00                | 1  | 2.03           | -13.96 | 4.05   |  | 1.43 |
| AB             | -9.570               | 1  | 2.28           | -15.16 | -3.99  |  | 1.63 |

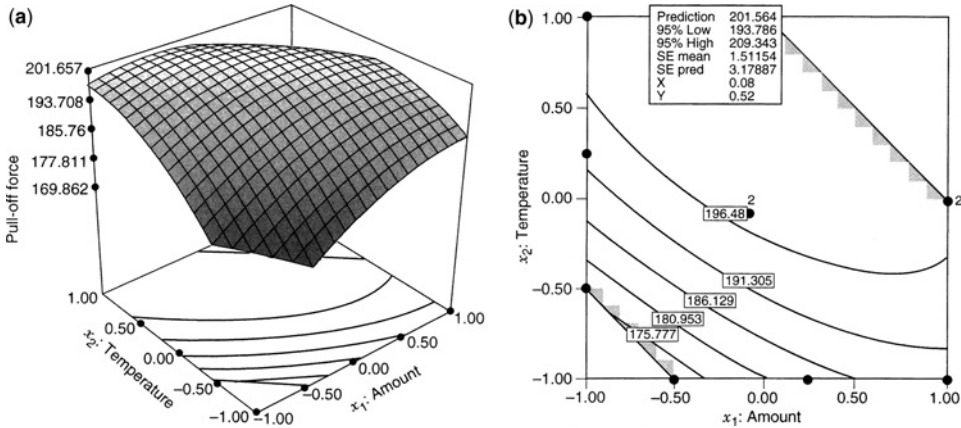
Final Equation in Terms of Coded Factors:

$$\begin{aligned}
 \text{Force} &= \\
 &+199.10 \\
 &+4.51 \quad *A \\
 &+9.56 \quad *B \\
 &-6.19 \quad *A^2 \\
 &-9.00 \quad *B^2 \\
 &-9.57 \quad *A*B
 \end{aligned}$$

“*Sequential Model Sum of Squares*”: Select the highest order polynomial where the additional terms are significant.

“*Lack of Fit Test*”: Want the selected model to have insignificant lack-of-fit.

“*Model Summary Statistics*”: Focus on the model maximizing the “Adjusted R-squared”; and the “Predicted R-Squared.”



**Figure 9.20** Plots of predicted pulloff force, Example 9.13. (a) Response surface. (b) Contour plot, showing location of the stationary point.

where  $z_3$  is an indicator variable that has a value of 0 for the first level of the qualitative factor and has a value of 1 for the second level. For this model, the assumption is that the shape of contour plots describing the relationship between  $x_1, x_2$  and the response is the same for both levels of the qualitative factor, except for a potential constant shift up or down between levels.

An alternate model might be

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{12}x_1x_2 + \delta_3z_3 + \delta_{13}x_1z_3 + \delta_{23}x_2z_3 + \epsilon$$

where the additional interaction terms now allow the main effects for  $x_1$  and  $x_2$  to be different for each level of the qualitative factor, but the interaction effect involving  $x_1$  and  $x_2$  is expected to be the same.

A final model that allows maximum flexibility would be

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{12}x_1x_2 + \delta_3z_3 + \delta_{13}x_1z_3 + \delta_{23}x_2z_3 + \delta_{123}x_1x_2z_3 + \epsilon$$

where the  $\delta_{123}x_1x_2z_3$  terms allows completely separate first-order with interaction models to be fit for each of the two levels of the qualitative factor. In general, if the experimenter does not know which of these models is most appropriate, it is helpful to assume the largest model when selecting a design, as this will allow the estimation and assessment of all terms in this model. However, protecting against the largest possible model will in general lead to larger designs.

When evaluating the prediction variance characteristics of a design, some adjustments are also needed for models including qualitative factors. For example, when the FDS plot is calculated, design locations cover the continuous range of all of the qualitative factors, but only examine the set levels of the qualitative factor(s). In this way, we can think of the design region as a series of separate subregions—one for each level of the qualitative factor. The possible design locations are selected throughout each subregion.

**Example 9.14** In an extraction experiment an engineer wishes to model the amount of extraction against temperature and time. The eventual goal may be to determine temperature and time values that maximize extraction. However, it is natural to consider a third factor: type of extraction solvent. One may view this qualitative factor as having three levels, because three potential candidate solvents exist, each with a distinctive chemical composition.

Computer-generated designs are a powerful tool when one encounters qualitative variables that must be modeled with the usual continuous quantitative design variables. Here, of course, the user must be concerned about the interaction that exists between continuous and qualitative factors. It is believed that the effect of temperature or time may change depending on which of the solvents is being considered.

Consider the situation in which we would like to model the amount of extraction as a function of type of solvent (with solvents A, B, and C) in addition to temperature and time, two standard continuous design variables. The model assumed is second order in  $x_1$  (temperature) and  $x_2$  (time) and the qualitative variable (solvent) interacts with  $x_1$  and  $x_2$ . The required design size is  $N = 15$ . Formally the model is written

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{12}x_1x_2 + \beta_{11}x_1^2 + \beta_{22}x_2^2 + \gamma_1z_1 + \gamma_2z_2 + \delta_{11}z_1x_1 + \delta_{21}z_2x_1 + \delta_{12}z_1x_2 + \delta_{22}z_2x_2 + \varepsilon$$

Here

$$z_1 = \begin{cases} 1 & \text{if A is the discrete level} \\ 0 & \text{elsewhere} \end{cases}$$

$$z_2 = \begin{cases} 1 & \text{if B is the discrete level} \\ 0 & \text{elsewhere} \end{cases}$$

Notice that indicator variables are used to represent the qualitative factor, type of solvent. The  $N = 15$  computer generated design using  $D$ -optimality is given by

$$D = \begin{matrix} & x_1 & x_2 & x_3 \\ \begin{bmatrix} -1 & -1 & C \\ -1 & -1 & B \\ -1 & -1 & A \\ -1 & 1 & C \\ -1 & 1 & B \\ -1 & 1 & A \\ 1 & -1 & C \\ 1 & -1 & B \\ 1 & -1 & A \\ 1 & 1 & C \\ 1 & 1 & B \\ 1 & 1 & A \\ 0 & 0 & A \\ 0 & 1 & B \\ -1 & 0 & B \end{bmatrix} \end{matrix}$$

**TABLE 9.7** Design Expert *D*-Optimal Design for Three Quantitative and Two Qualitative Variables

| Obs | $x_1$ | $x_2$ | $x_3$ | $z_1$ | $z_2$ |
|-----|-------|-------|-------|-------|-------|
| 1   | -1    | -1    | 1     | L1    | L1    |
| 2   | -0.5  | 0     | 0     | L1    | L1    |
| 3   | 1     | 1     | 1     | L1    | L1    |
| 4   | 1     | 1     | 1     | L1    | L1    |
| 5   | 1     | 1     | -1    | L1    | L1    |
| 6   | 1     | -1    | 1     | L1    | L1    |
| 7   | 1     | -1    | -1    | L1    | L1    |
| 8   | -1    | 1     | -0.33 | L1    | L2    |
| 9   | -1    | 0.33  | -1    | L1    | L2    |
| 10  | -1    | 0.33  | -1    | L1    | L2    |
| 11  | -1    | 0.33  | 1     | L1    | L2    |
| 12  | 0     | -1    | 0     | L1    | L2    |
| 13  | 0.33  | 1     | -1    | L1    | L2    |
| 14  | 1     | 0     | 0     | L1    | L2    |
| 15  | 1     | -1    | -1    | L1    | L2    |
| 16  | -1    | -1    | -1    | L1    | L2    |
| 17  | -1    | 1     | 1     | L2    | L1    |
| 18  | -1    | 1     | -0.33 | L2    | L1    |
| 19  | -1    | 0.33  | 1     | L2    | L1    |
| 20  | -0.33 | 1     | -1    | L2    | L1    |
| 21  | 0     | -1    | 0     | L2    | L1    |
| 22  | 1     | 1     | -1    | L2    | L1    |
| 23  | 1     | 1     | -1    | L2    | L1    |
| 24  | -1    | 1     | 0     | L2    | L2    |
| 25  | -1    | -1    | 0     | L2    | L2    |
| 26  | 0     | -1    | 1     | L2    | L2    |
| 27  | 0     | 0     | 0     | L2    | L2    |
| 28  | 1     | 1     | 0     | L2    | L2    |
| 29  | 1     | 0     | 1     | L2    | L2    |
| 30  | 1     | -1    | -0.33 | L2    | L2    |

Note the **crossing** of the  $2^2$  factorial points with the three levels *A*, *B*, and *C* in order to capture the interactions  $x_1z_1$ ,  $x_1z_2$ ,  $x_2z_1$ ,  $x_2z_2$ , and  $x_1x_2$ . The center and two axial points in  $x_1$  and  $x_2$  allow further information for estimation of  $x_1^2$  and  $x_2^2$  terms.

**Example 9.15 Designs for Three Quantitative and Two Qualitative Variables** Consider a situation in which there are three quantitative variables  $x_1$ ,  $x_2$ , and  $x_3$  and two qualitative variables  $z_1$  and  $z_2$ , each at two levels. Assume that the practitioner wishes to fit a full second-order model in the three quantitative variables as well as the terms  $z_1$ ,  $z_2$ ,  $z_1x_j$ ,  $z_2x_j$ ,  $z_1x_j^2$ , and  $z_2x_j^2$  for  $j = 1, 2, 3$ . There are 24 model terms to be fitted, including  $\beta_0$ . Using Design-Expert, a *D*-optimal design was created with 30 runs, specifying that the additional six degrees of freedom should be allocated with three degrees of freedom for lack of fit and three replicates. Table 9.7 shows the design.

The reader should note that in this example we are not assuming a separate response surface for all four combinations of the qualitative variables. We are in fact assuming that

**TABLE 9.8 JMP *I*-Optimal Design for Three Quantitative and Two Qualitative Variables**

| Obs | $x_1$ | $x_2$ | $x_3$ | $z_1$ | $z_2$ |
|-----|-------|-------|-------|-------|-------|
| 1   | -1    | 0     | 0     | L1    | L1    |
| 2   | -1    | -1    | 1     | L1    | L1    |
| 3   | 0     | 0     | -1    | L1    | L1    |
| 4   | 0     | -1    | 0     | L1    | L1    |
| 5   | 1     | -1    | -1    | L1    | L1    |
| 6   | 1     | 1     | -1    | L1    | L1    |
| 7   | 1     | 1     | 1     | L1    | L1    |
| 8   | -1    | 1     | 1     | L1    | L2    |
| 9   | -1    | 1     | -1    | L1    | L2    |
| 10  | 0     | 1     | 0     | L1    | L2    |
| 11  | 0     | 0     | 1     | L1    | L2    |
| 12  | 1     | 1     | -1    | L1    | L2    |
| 13  | 1     | 0     | 0     | L1    | L2    |
| 14  | 1     | -1    | 1     | L1    | L2    |
| 15  | -1    | -1    | -1    | L2    | L1    |
| 16  | -1    | 1     | 1     | L2    | L1    |
| 17  | 0     | 0     | 1     | L2    | L1    |
| 18  | 0     | 0     | 0     | L2    | L1    |
| 19  | 0     | -1    | 0     | L2    | L1    |
| 20  | 1     | -1    | 1     | L2    | L1    |
| 21  | 1     | 1     | -1    | L2    | L1    |
| 22  | 1     | 0     | 0     | L2    | L1    |
| 23  | -1    | -1    | 0     | L2    | L2    |
| 24  | -1    | 1     | -1    | L2    | L2    |
| 25  | 0     | -1    | -1    | L2    | L2    |
| 26  | 0     | 0     | 0     | L2    | L2    |
| 27  | 0     | 1     | 0     | L2    | L2    |
| 28  | 0     | 0     | 1     | L2    | L2    |
| 29  | 1     | 1     | 1     | L2    | L2    |
| 30  | 1     | 0     | 0     | L2    | L2    |

the  $x_i x_j$  ( $i \neq j$ ) terms are the same at the four combinations. Indeed, a model assuming separate response surfaces would require 16 additional model terms containing  $x_i x_j z_k$  terms ( $i \neq j$ ),  $z_1 z_2 x_i$  terms,  $z_1 z_2 x_i^2$  terms and  $z_1 z_2 x_i x_j$  terms ( $i \neq j$ ). Thus, in our model we assume stability of  $x_i x_j$  across  $z_1$  and  $z_2$  and no interaction among  $z_1$  and  $z_2$ .

Note some of the symmetry in the design: all four of the combinations of the qualitative variables have seven or eight observations, and the three replicates are spread with one pair in three of the four combinations.

Table 9.8 shows the *I*-optimal design generated by JMP, with no restrictions placed on the degrees of fit allocated for lack of fit or replication. This design also has seven or eight observations for each of the qualitative variables. Note how more of the continuous variable levels are selected with a value of 0 to give better prediction near the center and to lower the average prediction variance value. When no degrees of freedom are allocated for estimating lack of fit, then only three levels of each of the quantitative variables are selected. In addition, no replicates are selected if they are not requested.

Computer-generated designs are particularly useful for situations with a combination of qualitative and quantitative variables, as they provide great flexibility for accommodating the model structure required.

**Definitive Screen Designs with Two-level Qualitative Variables** As we saw in Section 8.4.8, Definitive Screening Designs can provide good screening designs for first-order with interaction models with the ability also estimate second-order effects. However, one of the limitations of the original DSDs is that they involve three levels for all of the factors. Jones and Nachtsheim (2013) describe two search-based algorithms to create DSDs suitable for including one or more qualitative factors with two levels. Both methods involve adding supplemental runs to the design that are chosen to have desirable properties for estimating the model of the form

$$y = \beta_0 + \sum \beta_i x_i + \sum \sum \beta_{ij} x_i x_j + \sum \delta_k z_k + \sum \sum \delta_{ik} x_i z_k + \sum \sum \delta_{kl} z_k z_l + \epsilon$$

where the  $x_i$ 's are the quantitative factors and the  $z_k$ 's are the qualitative factors. Note that the assumed model allows for all of the second-order interactions to be considered.

The first method, called the **DSD-augment** method, produces designs where all of the estimates of the main effects and second-order effects are independent, with some partial aliasing between two-factor interaction and the intercept. The designs are highly  $D$ -efficient. Below two such designs are shown for  $m = 4$  continuous factors with  $c = 1$  or 2 qualitative factors. The first four columns define the levels for the continuous variables, and the final 1 or 2 columns are for the qualitative factors.

|    | $m = 4, c = 1$ | $m = 4, c = 2$ |
|----|----------------|----------------|
| 1  | 0++++          | 0+++++         |
| 2  | 0-----         | 0-----         |
| 3  | +0+--          | +0+--+         |
| 4  | -0-++          | -0-++-         |
| 5  | ++0+-          | ++0+--         |
| 6  | --0-+          | --0-++         |
| 7  | +--0+          | +--0+-         |
| 8  | -+-0-          | -+-0-+         |
| 9  | +---++         | +---+++        |
| 10 | -++--          | -++----        |
| 11 | ++--+          | ++--++         |
| 12 | --++-          | --++--         |
| 13 | 0000+          | 0000--         |
| 14 | 0000-          | 0000++         |

The second method, called the **ORTH-augment** method, provides  $D$ -optimal designs. A disadvantage of this method is that there is partial aliasing between main effects and two-factor interactions for those terms involving the categorical factors. The corresponding designs for the scenarios listed above are given here.

|    | $m = 4, c = 1$ | $m = 4, c = 2$ |
|----|----------------|----------------|
| 1  | 0++++          | 0+++++         |
| 2  | 0-----         | 0-----         |
| 3  | +0+---         | +0+--+         |
| 4  | -0-++          | -0-++-         |
| 5  | ++0+-          | ++0+--         |
| 6  | --0-+          | --0-++         |
| 7  | +--+0+         | +--+0+-        |
| 8  | -+-0-          | -+-0-+         |
| 9  | +---++         | +-----         |
| 10 | -++++          | -++++-         |
| 11 | ++---+         | ++---++        |
| 12 | --++-          | --+++-         |
| 13 | 0000-          | 0000--         |
| 14 | 0000-          | 0000--         |
| 15 |                | 0000-+         |
| 16 |                | 0000+-         |

When comparing the results from the DSD-augment and Orth-augment methods, there are only subtle differences in the designs. For the  $m = 4, c = 1$  case, only runs 10 and 14 differ, with run 10 not being the negative complement of run 9, and both of the center runs being placed in the same level of the categorical factor. Similarly for the  $m = 4, c = 1$  case, only runs 10 and 12 differ, plus there is a need for two additional center runs. For designs from both methods for other scenarios involving 4 or 5 quantitative factors with 1 to 4 categorical factors, see Jones and Nachtsheim (2013).

## 9.4 MULTIPLE OBJECTIVE COMPUTER-GENERATED DESIGNS FOR RSM

In Section 7.2.4, Pareto front optimization was described for multiple responses. A similar approach can be used for finding an optimal design when considering several different design criteria. In Section 8.1, eleven desirable properties of design for consideration were described. They can be divided into several categories, one of which considers good estimation and prediction. In previous sections we have spent considerable time investigating designs that were generated using  $D$ -,  $G$ -, and  $I$ -optimality as the criterion for computer-generated designs. Recall that one of the cautions against using computer-generated designs was their exclusive focus on a single objective, which often sacrificed performance in other areas that the Central Composite, Box–Behnken and Definitive Screening Designs naturally accounted for.

In this section, we describe an approach that allows the primary objectives captured in the alphabetic optimality criteria to be considered in conjunction with other aspects of good design. Other aspects that might be natural to try to include are the ability to investigate lack of fit or protect against bias introduced from model misspecification. This allows the experimenter to cope with model uncertainty, while still pursuing a design that performs well for the assumed model. Another aspect that might be considered as a complementary criterion is the ability to estimate pure error. All of the tests of hypothesis about the significance of model terms are based on having a good estimate of natural variation, and

the use of replicates allows for an estimate of this that is independent of the model specified. Lastly, there may be some flexibility in the size of the design to be run—and looking at the trade-offs between cost and the improvements in estimation and prediction is an effective way to determine what size of experiment makes the most sense. However, the approach is general and can be applied to any set of quantitative measures that summarize the desirable aspects of a designed experiment that the user wishes to formally incorporate into design construction or augmentation.

#### 9.4.1 Pareto Front Optimization for Selecting a Design

When using a single criterion for constructing or augmenting an optimal design, a single solution is typically identified at the conclusion of the search. While this is comfortably straightforward to determine how to proceed, it does mask the fact that there often might be other contenders that were near-optimal, which might have provided practically the same performance had they been used. The Pareto front approach seeks to provide a set of alternatives that can be considered in more detail to determine what the best choice is for a particular scenario. Most often, we acknowledge that the selection of a best design should be based on multiple criteria that characterize the “goodness” of the design. Hence, the Pareto front approach with its identification of a set of solutions provides context for examining alternatives and deciding on their relative merits.

In Section 7.3, the DMRCs process for structured decision-making was described which involved two general stages. In the first, the set of all solutions (those located on the Pareto front) that are objectively best are identified, and they are said to dominate all of the other possible designs. Hence designs not on the Pareto front can be discarded from further consideration, as long as the user feels confident that they have identified the correct set of criteria on which to focus. The second stage is subjective and consists of examining the alternatives on the Pareto front identified in the first stage to choose the best one that balances the criteria in a way best suited to the goals of the experiment. In this stage, graphical methods are commonly used to help visualize the tradeoffs between alternatives. The Pareto front approach is flexible to handle different numbers of criteria, with a general rule of thumb of restricting the optimization to a moderate number (2–4) of objectives. If too many criteria are simultaneously considered, then the number of possible solutions can become unmanageable, and it becomes difficult to balance too many trade-offs.

In many data collection exercises using designed experiments, the goal is to provide information for several engineers or several purposes. Hence allowing for different user priorities to be considered can lead to improved choices for the best design to select. There are several key differences between using the Pareto front approach for responses described in Chapter 7 and when the methods are used for the construction and augmentation of designed experiments:

1. Most responses are naturally quantitative. For designed experiments, there may be an informal sense of wanting a design to have a general characteristic, but there may not be an automatic quantitative measure that summarizes this. Hence, it may require some work to identify metrics for a particular characteristic, such as the ability to assess lack of fit or protect against bias, that appropriately reflect the priorities of the study.
2. The metrics for a design should not have uncertainty associated with them. Recall for responses that were estimated from observed data, there was uncertainty associated



with what values might be obtained for new observations. This needed to be incorporated into the optimization strategy to provide realistic summaries about which combination of factor values might lead to an optimal solution. For the designed experiment assessment and optimization, each design will have a single value of the selected metric to be optimized, and so incorporating uncertainty is not needed. This is a helpful simplification.

3. In Chapter 7, the optimization was based on discretizing the design locations into a grid of possible choices, and then the Pareto front was constructed from the set of alternatives. For multiple objective optimization of designs, there is rarely an itemized list of choices from which to select. Hence, a search algorithm is needed to efficiently explore the vast number of possible designs to find the Pareto front of desirable solutions. Section 9.4.2 describes one algorithm that can be used to find the set of designs on the Pareto front.
4. With responses modeled using low-order polynomials, the surfaces summarizing the relationship between inputs and responses are continuous and smooth. This is helpful as adjacent points in the design space share similar predicted response values. Hence when optimizing responses, as in Chapter 7, the key is to identify regions of good performance. When we consider optimizing designs, we find that changing a single run can dramatically alter the characteristics of the design. This is because the runs in a design are very much interdependent and estimating the model parameters relies on information from multiple observations. Hence the search for an optimal design will need to take into account the potentially very different characteristics of similar designs. Nearly identical designs can perform very differently for some criteria, and very different designs can sometimes give similar performance across the criteria of interest.

As noted in item 3 above, an efficient search algorithm is a key part of the success of using the DMRCs process for selecting a design based on several objectives. In the next section, we describe a possible algorithm for finding the Pareto front of nondominated solutions.

#### 9.4.2 Pareto Aggregating Point Exchange Algorithm

The set of possible designs for even moderately small design problems can become large very quickly. As a simple illustration, imagine that an experimenter wishes to consider a 12-run design for an experiment involving three factors. Since there is interest in fitting a second-order model in a cuboidal region, each factor is restricted to just 3 levels (-1, 0, +1 in coded variables). Then there are  $3^3 = 27$  candidate points for each run, as well as 12 runs. The total number of possible designs then is  $27^{12} \approx 1.5 \times 10^{17}$ . If the assumed model is a first-order model and we consider two levels with  $2^3 = 8$  candidate points, there are still  $8^{12} \approx 6.9 \times 10^{10}$  possible designs. It should be noted that many of these design possibilities would be nonsensical choices—for example, choosing all 12 runs to be the same candidate point—but the search algorithm still needs to navigate and evaluate a vast set of options.

The Pareto Aggregating Point Exchange (PAPE) algorithm [Lu, Anderson-Cook, and Robinson (2011) and Lu and Anderson-Cook (2013)] is an adaptation of point exchange algorithms described in Section 9.3.2, with several key differences. First, the common aspects are that the algorithm requires the user to specify the same aspects of the design as before: the assumed model, the required sample size, the ranges on the design variables,

the candidate points, and any additional constraints. In addition, the algorithm also requires the user to specify two or more quantitative metrics used to evaluate the designs.

One important difference between this as a standard point exchange algorithm is the overall goal of the algorithm. When optimizing is based on a single criterion, the final result is a single solution. With this framework, it is easy to evaluate a new candidate solution: Does it perform better than the current best choice? If it is better, then we keep this value and discard the previous best solution. If it is not better, then we keep the current solution and move on to evaluate another candidate.

However, with the PAPE algorithm (or any multiple objective optimization), the goal is to construct the Pareto front of nondominated solutions. Hence the result of the algorithm is not a single solution, but a set of solutions. In the point exchange algorithm, updates occur when the current candidate point for different runs of the design are compared to the other available candidates to look for improvements. But what does improvement mean here? What mechanism should be used to decide if we keep the current run or update to a new candidate? The PAPE algorithm uses a grid of fixed weights from a desirability function (as described in Section 7.2.3) to define different directions for the search algorithm. For example, suppose that an experimenter wishes to create a Pareto front of designs that balance optimizing  $D$ - and  $I$ -optimality. Then the PAPE algorithm might run six parallel searches from each random start: The first optimizes based on only  $D$ -efficiency, the second optimizes based on  $0.8*D$ -efficiency +  $0.2*I$ -efficiency. The remaining searches would use  $(0.6,0.4)$ ,  $(0.4,0.6)$ ,  $(0.2,0.8)$ ,  $(0,1)$  as the weights for the two criteria. This allows for quick decisions based on a single number from the desirability function about whether a new design is better or worse than the current candidate, and it spreads out the directions of the search. Lu, Anderson-Cook, and Lin (2014) describe an approach for looking at a more focused set of weights if the experimenter knows a priori how to prioritize the different criteria.

The second major difference between a standard point exchange algorithm and the PAPE algorithm results from wanting a set of solutions instead of a single choice. As different candidate points are evaluated as part of the search, each new design is compared to the existing Pareto front to determine if it should be included. This means that for a single random start, many possible designs could potentially be added to the front. Since the calculation of all of the criteria has already been done as part of the search, this can dramatically speed how quickly the final Pareto front is populated.

The algorithm can be adapted to different numbers and types of quantitative design criteria. While this optimization typically takes considerably longer than optimizing over a single objective, it is still relatively short compared to the time, effort, and cost of the actual running of the experiment.

### 9.4.3 Using DMRCS for Design Optimization

When constructing or augmenting a design using multiple objectives, the experimenter has many choices about what to prioritize as important for the design criteria. Some potential criteria used as one of the objectives for optimizing multiple facets of a good design include:

#### Assuming the chosen model is correct

- (a)  $D$ -efficiency—maximize for good parameter estimation
- (b)  $I$ -efficiency—maximize for good average prediction variance
- (c)  $G$ -efficiency—maximize for good worst case prediction variance

**Wanting some protection in case the chosen model is incorrect**

- (d)  $\text{tr}(\mathbf{AA}')$ —minimize to protect model terms from bias ( $\text{tr}(\mathbf{AA}')$  where  $A = (\mathbf{X}'_1\mathbf{X}_1)^{-1}(\mathbf{X}'_1\mathbf{X}_2)$  was introduced in Section 8.2.2)
- (e)  $\text{tr}(\mathbf{R}'\mathbf{R})$ —minimize to protect SSE (sum of squares error) from bias ( $\text{tr}(\mathbf{R}'\mathbf{R})$  where  $R = X_1A - X_2$  was introduced in Section 8.2.2)

**Estimating pure error with more degrees of freedom**

- (f) # of replicates

**Estimating lack-of-fit terms**

- (g)  $\text{tr}(\mathbf{R}'\mathbf{R})$ —maximize to provide estimates of lack of fit

**Cost**

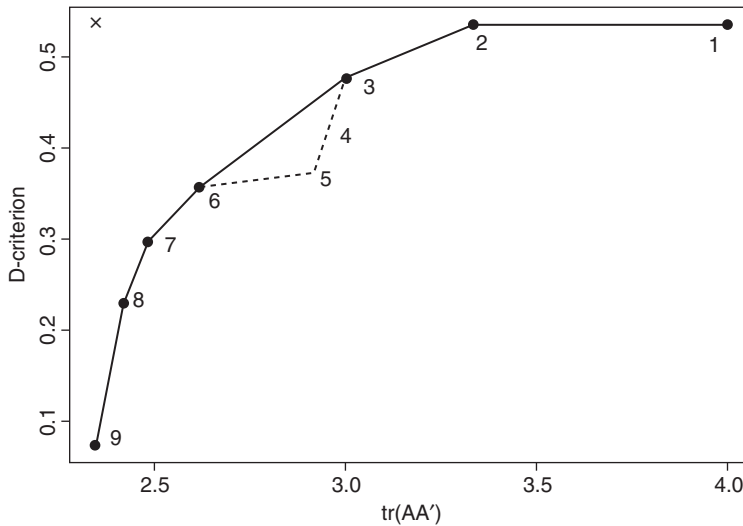
- (h) total # of runs

**Example 9.16 A Nonstandard Five-Factor Experiment** For a very expensive screening experiment, an engineer has sufficient funds to run a 14-run experiment with five-factors (A–E). Based on expertise from some colleagues, the proposed model to fit is a specialized first-order model with interactions of the form

$$Y = \beta_0 + \beta_A A + \beta_B B + \beta_C C + \beta_D D + \beta_E E + \beta_{AB} AB + \beta_{AC} AC + \beta_{BD} BD + \beta_{CE} CE + \varepsilon \tag{9.26}$$

where the interactions selected represent current understanding of the underlying process. Since the experiment must be necessarily be restricted to 14 runs, it is not possible to simultaneously estimate all of the 10 two-factor interactions if they had been included. Hence the goal is to construct a design using two levels for each of the factors (–1,+1) that allows good estimation of the model. A natural choice of metric for this criterion is to look at *D*-efficiency and try to maximize  $|M| = |\mathbf{X}'\mathbf{X}|/N^p$ . However, the engineer is concerned that at least one of the excluded two-factor interactions might be active, and she wishes to include a metric that considers the bias that may be introduced into the estimation of the model parameters in Equation 9.26. One option to balance *D*-efficiency with protection against bias is to try to minimize  $\text{tr}(\mathbf{AA}')$  where  $A = (\mathbf{X}'_1\mathbf{X}_1)^{-1}(\mathbf{X}'_1\mathbf{X}_2)$  and  $\mathbf{X}_1$  is the  $14 \times 10$  matrix with a column for the intercept, the five main effects, and the four selected two-way interactions (*AB*, *AC*, *BD*, and *CE*).  $\mathbf{X}_2$  is the  $14 \times 6$  matrix with a column for the omitted two-factor interactions (*AD*, *AE*, *BC*, *BE*, *CD*, and *DE*).

The PAPE algorithm is used to find the Pareto front with nine nondominated designs. Figure 9.21 shows a scatterplot of the values of  $|\mathbf{X}'\mathbf{X}|/N^p$  and  $\text{tr}(\mathbf{AA}')$ , with the *D*-optimal design shown in the top right corner of the plot. The design with the smallest value of  $\text{tr}(\mathbf{AA}')$  is in the bottom left. The Utopia point (the combination of the best  $|\mathbf{X}'\mathbf{X}|/N^p$  and  $\text{tr}(\mathbf{AA}')$ ) is shown in the top left corner, but there is no design that is able to achieve this combination of values. From the plot, it is possible to gain some understanding about the trade-offs between the two criteria. If *D*-optimality is the only objective, then the best design (labeled 1) has the potential to introduce considerable bias into the model parameter estimates if the model has been incorrectly specified. At the other extreme, if the design that minimizes  $\text{tr}(\mathbf{AA}')$  is selected (labeled 9), then the engineer will have larger associated variances for



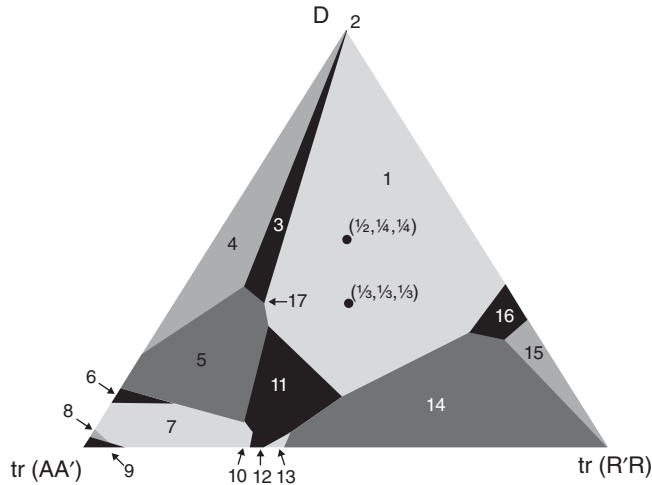
**Figure 9.21** Pareto front for Example 9.16 with nine nondominated designs. The designs on the convex hull of the Pareto front are shown with a solid line. The Utopia point is marked with an “x”.

the model parameters if the model is correct. Either extreme seems undesirable, but before the experiment is run, it is unknown whether the assumed model is correct or not.

The Pareto front allows for several alternatives that provide better balance between the extremes of the optimizing just a single objective. Clearly, design 2 feels like a superior choice to design 1, since for virtually no change in  $|\mathbf{X}'\mathbf{X}|/N^p$ , the potential bias can be reduce. This is again a reminder of the danger of blindly looking for an optimal design based on a single criterion.

The designs labeled 3 or 6 may be suitable choices for the experiment, since they provide good estimation of the model parameters while still reducing the amount of potential bias if the model is incorrectly specified. Designs 4 and 5 seem less appealing than either 3 or 6, since they are not on the convex hull of the Pareto front, and hence it feels like there is more sacrifice from one criterion relative to the improvement realized by the other. Which of the designs makes the most sense for the engineer depends on the relative prioritization that she wishes to make about the importance of the two criteria: How likely is the model to be correct? What are the consequences of having bias introduced into the model parameters? The graphical tools such as the best solutions for different weights in a desirability function (Fig. 7.3) and the trade-off plot (Fig. 7.4) can provide more quantitative tools for comparing alternatives. Recall that there is no single universal best answer, but rather the goal of using the Pareto front approach is to gain perspective on what alternatives are available and to determine which one is the best fit for study goals.

**Example 9.17 Adding a Third Criterion to Design Optimization** We now reconsider the same experiment described in Example 9.16, but include an additional criterion (Lu, Anderson-Cook, and Robinson, 2011). Previously the goal was to seek large values of  $|\mathbf{M}| = |\mathbf{X}'\mathbf{X}|/N^p$  for good  $D$ -efficiency and small values of  $\text{tr}(\mathbf{A}\mathbf{A}')$  to reduce the bias in the parameter estimates of the model parameters. Now in addition to these two criteria, we add the goal of minimizing the bias in estimating the natural variability of the process.

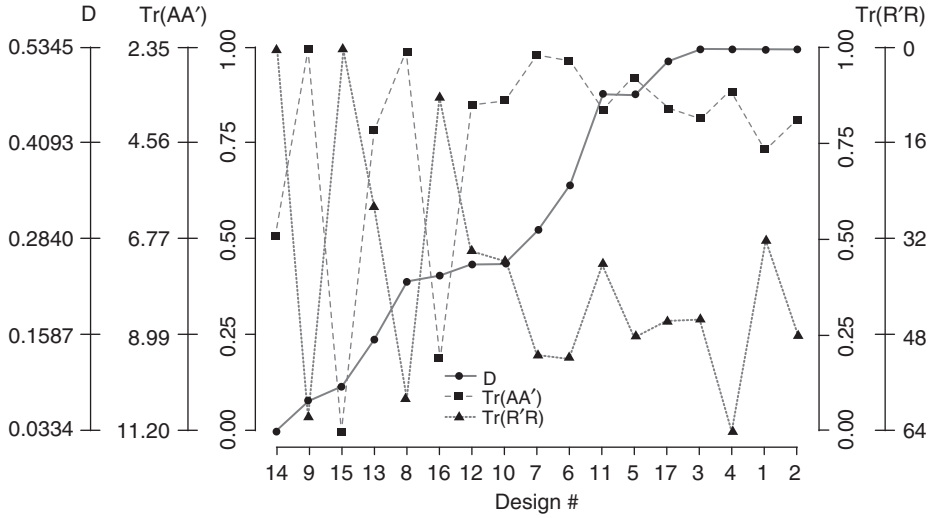


**Figure 9.22** Mixture plot of designs from Example 9.17 with criteria  $D$ -efficiency,  $\text{tr}(\mathbf{A}\mathbf{A}')$  and  $\text{tr}(\mathbf{R}'\mathbf{R})$ .

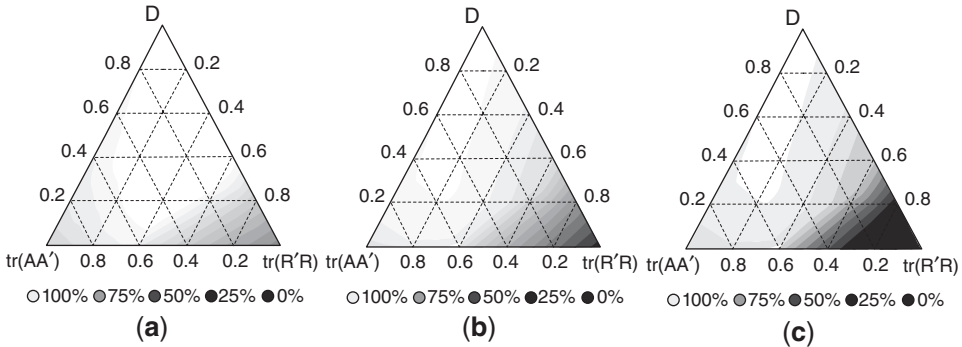
Having a good estimate of the error term variability helps with the calibration of test statistics to assess the terms included in the model. One potential choice of quantitative metric for this objective is to find models with small values of  $\text{tr}(\mathbf{R}'\mathbf{R})$ , where  $\mathbf{R} = \mathbf{X}_1\mathbf{A} - \mathbf{X}_2$  and  $\mathbf{A} = (\mathbf{X}'_1\mathbf{X}_1)^{-1}(\mathbf{X}'_1\mathbf{X}_2)$ . The  $\mathbf{X}$ -matrices are defined as before in Example 9.16.

When the Pareto front is found using the PAPE algorithm for the three criteria, there are 351 nondominated designs. Since this is an unmanageable number of designs to try to understand, a mixture plot similar to Fig. 7.13 is created to identify designs that are best for some combination of desirability function weights using an additive desirability function with scaling for the individual desirability scores based on the best and worst values for each criterion on the Pareto front. Figure 9.22 shows the mixture plot with 17 designs that are best for at least one combination of weights. (The numbers for the designs are not the same as the numbering used for Example 9.16.) Focusing on only those designs which are best for some weight combination reduces the number of designs to consider to a more manageable number. Notice how the design labeled 1 in Fig. 9.21 is best for a large proportion of the weight combinations, and it is best for several potentially natural weight combinations: The centroid of the mixture space corresponds to choosing weights of  $(w_D, w_{\text{tr}(\mathbf{A}\mathbf{A}')}, w_{\text{tr}(\mathbf{R}'\mathbf{R})}) = (1/3, 1/3, 1/3)$  for each of the criteria. If the engineer wanted to give equal weight to assuming the model was correct and protecting against bias, one weight combination that reflects that is  $(1/2, 1/4, 1/4)$ .

Figure 9.23 shows the trade-off plot of these designs, sorted from highest to lowest  $D$ -efficiency. The design labeled 1 is notable in the trade-off plot as the only solution that has an individual desirability score at or above 0.5 for all three criteria. In selecting where to place priorities between the three criteria, the engineer decides that it makes sense to place more weight on assuming that the model is correct and therefore weights  $D$ -efficiency higher. Using synthesized efficiency plots, similar to those introduced in Fig. 7.15, she examines the designs labeled 1, 3, and 4 in Figure 9.24. Given the high efficiency (white and light shades of gray) shown for most regions of the weight plot for the design labeled 1, she selects this as the best design for the intended experiment. Table 9.9 lists the 14 runs in the chosen design.



**Figure 9.23** Trade-off plot of designs from Example 9.17 with criteria  $D$ -efficiency,  $\text{tr}(\mathbf{A}\mathbf{A}')$  and  $\text{tr}(\mathbf{R}'\mathbf{R})$ .



**Figure 9.24** Synthesized efficiency plots of leading design candidates (a) design 1, (b) design 3, and (c) design 4 from Example 9.17 based on criteria  $D$ -efficiency,  $\text{tr}(\mathbf{A}\mathbf{A}')$  and  $\text{tr}(\mathbf{R}'\mathbf{R})$ .

**TABLE 9.9** Selected Best Design (Labeled Design 1 in Figs. 9.22 and 9.23) Using Pareto Front Optimization Based on  $D$ -Efficiency,  $\text{tr}(\mathbf{A}\mathbf{A}')$  and  $\text{tr}(\mathbf{R}'\mathbf{R})$  for Example 9.17

| Run | (A, B, C, D, E)      | Run | (A, B, C, D, E)      |
|-----|----------------------|-----|----------------------|
| 1   | (-1, -1, +1, -1, -1) | 8   | (+1, -1, -1, -1, +1) |
| 2   | (-1, -1, -1, -1, -1) | 9   | (+1, +1, -1, -1, +1) |
| 3   | (-1, +1, -1, -1, -1) | 10  | (+1, -1, +1, -1, +1) |
| 4   | (+1, -1, -1, +1, -1) | 11  | (+1, +1, +1, -1, +1) |
| 5   | (+1, -1, +1, +1, -1) | 12  | (-1, -1, -1, +1, +1) |
| 6   | (+1, +1, -1, +1, -1) | 13  | (-1, -1, +1, +1, +1) |
| 7   | (+1, +1, +1, +1, -1) | 14  | (-1, +1, +1, +1, +1) |

## 9.5 SOME FINAL COMMENTS CONCERNING DESIGN OPTIMALITY AND COMPUTER-GENERATED DESIGN

As we indicated earlier in this chapter, the current status of computer-generated design began as a result of the theoretical work of Keifer and coworkers. This foundation was extended with decades of work by many who made the concept usable by practitioners. There is no question that computer-aided design of experiments can be of value. However, in RSM work there are several reasons to proceed with caution. Those who successfully use design computer packages to aid in constructing RSM designs are those who do not use it as a “black box.” Computer-aided designs can be useful if the user allows it to complement the available arsenal of standard RSM designs. The user who is likely to fail with computer-aided designs is the one who thinks too narrowly about what should be considered as a good design and is not armed with important knowledge about design optimality and standard RSM designs.

The development of optimal design theory is based on the use of a continuous design measure that determines the proportion of runs that should be made at a number of points in a predetermined design space. In practice, of course, the concept of continuous designs is impractical because each run must be uniquely specified. However, the idea is to determine discrete designs that are close approximations of the optimal continuous designs. The ideas of candidate lists and sequential formation of the design according to the important augmentation relationship of Section 9.3.1 provide computational ease and make computer-aided design practical. However, perhaps because of this useful relationship,  $D$ -optimality or, rather, the maximization of  $D$ -efficiency remains a commonly used criterion. One must keep in mind that the  $D$ -criterion is necessarily quite narrow, based on crucial and often oversimplified assumptions, and does not address many of the 11 important goals of RSM designs discussed in Chapter 8. It focuses solely on the good estimation of model parameters for a single model that is assumed to be correct.

With the development of improved algorithms for finding  $I$ - and  $G$ -optimality, there are more choices for criteria based on good properties of prediction variance. As we saw in Section 9.3.3,  $I$ -optimality with its focus on the average prediction variance tends to yield overall better characteristics for the prediction variance throughout the entire design space. Hence, in most situations where precise estimation of new observations or optimization of the response(s) is important,  $I$ -optimality is a good choice.

Many of the remaining 11 characteristics of good designs address the notion of design robustness—that is, robustness to model misspecification, robustness to errors in control, and so on. The following are some items that should be a reminder for caution among practitioners in the use of computer-generated designs:

1. For  $D$ -,  $I$ -, and  $G$ -optimality, the user must indicate a model on which to base the optimization. Typically a large model with terms that are thought to potentially exist is specified, since the model is not known ahead of time, and model editing or reduction is part of any modeling strategy. If the fitted model is not that which is input to the computer for the optimization, the computer-generated design is no longer optimal, and, indeed, it may not be a good design at all. Heredia-Langner et al. (2004) present strategies for generating designs that are robust to model choice, while Ozol-Godfrey et al. (2005) present graphical tools to evaluate prediction variance for different edited models.

2. Designs generated from the  $D$ -criterion do not address prediction variance. In some cases, a  $D$ -efficient design may not have ideal properties for prediction. Similarly, an  $I$ -optimal design does not focus on good prediction of the model parameters, and hence may not have ideal properties for estimation.
3. Optimization based on a single criterion is by necessity a simplification of the goal of estimation. The  $D$ -optimality criterion  $|\mathbf{M}| = |\mathbf{X}'\mathbf{X}|/N^p$  considers a single number summary of good model parameter estimation combining all terms in the model. Similarly,  $I$ -optimality looks at just a global summary of a distribution of prediction variance values across a multi-dimensional design space. While search algorithms benefit from focusing on single number summaries, the experimenter should look more closely at the details of what characteristics a design delivers. Graphical methods such as the Variance Dispersion Graph, Fraction of Design Space plots, and the Prediction Variance profiler in JMP can provide the experimenter with richer understanding of the design's performance.
4. One strategy for increasing the robustness of the design is to begin with a  $D$ - or  $I$ -optimal design and then augment them with additional runs targeting additional desirable characteristics of good designs. Including replicates or center runs can help improve overall performance, and both JMP and Design-Expert provide an option for adding additional runs after an optimal design has been created.
5. Another strategy for building robust designs is to begin with several quantitative design criteria and use the Pareto front approach to find the set of all nondominated designs. Having a collection of good designs with different trade-offs between the formal criteria shows the range of performance possible for each objective, as well as the degree of trade-off needed to achieve good performance. By using graphical tools and different desirability function weightings, most promising designs can be highlighted for deeper investigation. Choosing a best design is made easier by having several promising candidates to compare, and the Pareto front allows experimenters' priorities to be included in the decision-making process. The Pareto front approach is flexible for the number and type of characteristics that can be considered simultaneously.

In RSM analyses, the natural starting point for selecting a design should be to search for an appropriate standard design discussed in Chapters 8 and 9. Nevertheless, there will be circumstances under which a computer-generated design will be helpful. Constraints regarding sample sizes, the use of qualitative variables, and cases where the user must deal with an unusual combination of ranges or constraints are situations where computer-aided design may be helpful. Augmenting computer-generated designs with additional runs can help make designs perform better in many of the 11 criteria listed in Chapter 8. Drain et al. (2004) and Lin et al. (2015) present general methodology for constructing optimal RSM designs using a genetic algorithm. However, the user should be cautious regarding focusing too narrowly on a single objective, as well as with regard to being unduly confident in the assumed model.

Performing an experiment generally has multiple goals and associated uncertainty about the underlying assumptions of the model. The process used to select the appropriate design for the experiment should acknowledge the complexity of the goals and assumptions and should incorporate multiple facets of a good design into the selection.



## EXERCISES

- 9.1** There is considerable interest in the use of saturated or near saturated second-order designs. One interesting comparison involves the hexagon ( $k=2$ ) and the small composite design. Both involve six design points (plus center runs). In this exercise, we make comparisons that take into account variances of coefficients and prediction variance. However, before any two designs can be compared, they must be adjusted so they have the same scale. Consider the hexagon and the  $k=2$  small composite design. For proper comparisons we have

$$\mathbf{D}_{\text{SCD}} = \begin{array}{c} x_1 \quad x_2 \\ \left[ \begin{array}{cc} -1 & -1 \\ 1 & 1 \\ -\sqrt{2} & 0 \\ \sqrt{2} & 0 \\ 0 & -\sqrt{2} \\ 0 & \sqrt{2} \\ 0 & 0 \\ 0 & 0 \end{array} \right] \end{array} \quad \mathbf{D}_{\text{HEX}} = \begin{array}{c} x_1 \quad x_2 \\ \left[ \begin{array}{cc} \sqrt{2} & 0 \\ \sqrt{2}/2 & \sqrt{3}/2 \\ -\sqrt{2}/2 & \sqrt{3}/2 \\ \sqrt{2} & 0 \\ -\sqrt{2}/2 & -\sqrt{3}/2 \\ \sqrt{2}/2 & -\sqrt{3}/2 \\ 0 & 0 \\ 0 & 0 \end{array} \right] \end{array}$$

For this scaling, both designs are on spheres of radius  $\sqrt{2}$  (apart from center runs).

- (a) Compute  $N \text{Var}(b_0)/\sigma^2$ ,  $N \text{Var}(b_{ii})/\sigma^2$ ,  $N \text{Var}(b_i)/\sigma^2$ ,  $N \text{Var}(b_{ij})/\sigma^2$  for both designs. Comment.
- (b) Compute  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  at the design center and at all design points for both designs. Is this an illustration of rotatability?
- (c) Construct FDS plots and VDGs for both designs. Comment on which design you prefer and why.
- (d) From part (b), is there an illustration of the rotatability of the hexagon?
- (e) Is the small composite design rotatable? Explain.
- 9.2** Consider a comparison between the  $k=4$  SCD of Equation 9.5 with the  $k=4$  CCD ( $N=24+n_c$ ). Use  $n_c=4$  center runs for each design, and use  $\alpha=2$  for both designs.
- (a) Do the designs need to be rescaled? Explain.
- (b) Make the same type of comparisons as in Exercise 9.1.
- 9.3** Perhaps you noticed some evidence in Exercise 9.1 that the  $k=2$  SCD was not rotatable. Actually the conditions required for rotatability cannot be met for any SCD that involves a resolution III\* fraction of a  $2^k$ . Prove this result. *Hint*: Look at odd moments.
- 9.4** Create FDS and VDG plots for the Hoke  $\mathbf{D}_2$  and  $\mathbf{D}_6$  designs for  $k=3$ . Compare their prediction variance to the CCD with two center runs on a cuboidal region. Which design seems preferable? Explain.
- 9.5** The Koshal design is an interesting experimental plan that was not derived with statistical analysis in mind. Analysts are often surprised at the applicability and the

efficiency of the design, particularly when information regarding interaction is either known or considered to be unimportant. The Koshal design is a natural competitor of the saturated or near-saturated SCD. Compare the  $k = 3$  Koshal design with the  $k = 3$  ( $N = 10 + n_c$ ) SCD. Again, use scaled variances of regression coefficients. In addition, use  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  at the following locations in the *scaled design units*:

- (0, 0, 0)
- (1, 0, 0)
- (-1, 0, 0)
- (1.5, 0, 0)
- (-1.5, 0, 0)
- (1, 1, 1)
- (-1, -1, -1)

The design scaling should be as follows:

$$\mathbf{D}_{\text{SCD}} = \begin{matrix} & x_1 & x_2 & x_3 \\ \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \\ -\sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & 0 \\ 0 & -\sqrt{3} & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & -\sqrt{3} \\ 0 & 0 & \sqrt{3} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & & & \end{matrix} \qquad \mathbf{D}_{\text{Kosh}} = \begin{matrix} & x_1 & x_2 & x_3 \\ \begin{bmatrix} \sqrt{3/2} & 0 & 0 \\ 0 & \sqrt{3/2} & 0 \\ 0 & 0 & \sqrt{3/2} \\ -\sqrt{3/2} & 0 & 0 \\ 0 & -\sqrt{3/2} & 0 \\ 0 & 0 & -\sqrt{3/2} \\ \sqrt{3/2} & \sqrt{3/2} & 0 \\ \sqrt{3/2} & 0 & \sqrt{3/2} \\ 0 & \sqrt{3/2} & \sqrt{3/2} \end{bmatrix} & & & \end{matrix}$$

Comment on your results.

- 9.6 Compare, in terms of scaled regression coefficients, the  $k = 3$  311B (use  $n_c = 3$ ) with the CCD and SCD values in Table 9.1. The 311B must, of course, be scaled properly. Thus, divide it by  $\sqrt{2}$ . Comment on your results.
- 9.7 Consider the following first-order design

$$\mathbf{D} = \begin{matrix} & x_1 & x_2 & x_3 \\ \begin{bmatrix} -1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & 1 & 1 \\ -1 & -1 & 1 \end{bmatrix} & & & \end{matrix}$$

- (a) Is the design first-order orthogonal? Explain?  
 (b) Give the  $D$ -efficiency of this design.  
 (c) Give the  $I$ -efficiency of the design.  
 (d) Give the  $A$ -efficiency of the design. *Hint:* The  $A$ -efficiency of a design is given by

$$\frac{\text{Min}_{\zeta} \text{tr}[\mathbf{M}(\zeta)]^{-1}}{N \text{tr}(\mathbf{X}'\mathbf{X})^{-1}}$$

Now, of course, for a first-order model with range  $[-1, +1]$  on each design variable we have

$$\text{Min}_{\zeta} \text{tr}[\mathbf{M}(\zeta)]^{-1} = p$$

- 9.8** Consider a model containing first-order and two-factor interaction terms in three design variables—that is,

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3$$

In Section 9.2.1 we demonstrated that a resolution III two-level design is  $D$ -,  $A$ -,  $G$ -, and  $I$ -optimal for a first-order model. The levels must be at  $\pm 1$  extremes.

- (a) Show that all variances of coefficients are minimized with a resolution IV two-level design with levels at  $\pm 1$  extremes.  
 (b) Using the result in (a), show that the members of the same resolution IV class are  $I$ -optimal.  
 (c) Show that these designs are also  $G$ -optimal. The region  $R$  is, of course, the unit cube. Make use of the result (Eq. 9.14)

$$\text{Max}_{\mathbf{x} \in R} [v(\mathbf{x})] \geq p$$

and show that, for the class of designs in question,

$$\text{Max}_{\mathbf{x} \in R} [v(\mathbf{x})] = p$$

where, of course,  $p = 7$ .

- 9.9** Consider the Hoke  $\mathbf{D}_2$  design for  $k = 4$  assuming a second order model.  
 (a) Calculate the  $\mathbf{X}'\mathbf{X}$  matrix for this design.  
 (b) Is the design rotatable? Explain.  
 (c) Generate an FDS plot for the design.  
 (d) Is the design  $G$ -optimal? If not, what is its  $G$ -efficiency?

9.10 Consider the following Koshal design for  $k = 3$ :

$$\mathbf{D} = \begin{array}{c} x_1 \quad x_2 \quad x_3 \\ \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \end{array}$$

Assume that the region of design operability is  $[-1, +1]$  on each design variable. Consider the model

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3$$

- (a) Compute the  $D$ -efficiency for this design for the above model.  
 (b) Compute the  $A$ -efficiency.  
 (c) Compute the  $G$ -efficiency.
- 9.11 Often a standard second-order design is planned because the practitioner expects that the fitted model will, indeed, be second-order. However, when the analysis is conducted, it is determined that the model is considerably less than order 2. As an example, suppose a  $k = 2$  CCD with  $\alpha = 1.0$  and two center runs is used and all second-order effects (quadratic and interaction) are very insignificant. The design is used eventually for a first-order model.
- (a) What is the  $D$ -efficiency for a  $k = 2$  CCD with  $\alpha = 1.0$  and two center runs for a first-order model? Comment.  
 (b) How does the  $D$ -efficiency change if there are no center runs?  
 (c) Is the difference between the results in (a) and (b) expected?  
 (d) Give the  $A$ -efficiency for both designs.
- 9.12 Refer to Exercise 9.11. Give the  $D$ -efficiency of a  $k = 3$  BBD for the model

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3$$

9.13 Consider the following experimental design:

$$\mathbf{D} = \begin{array}{c} x_1 \quad x_2 \quad x_3 \\ \begin{bmatrix} -1 & -1 & -1 \\ 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

A second stage is used, as it becomes obvious that the model is quadratic. Use a computer-generated design package to augment the above design with six more design points in order to accommodate a complete second-order model.

- 9.14** Consider a first-order orthogonal design with  $\pm 1$  levels. Show that the addition of center runs *must* lower the  $D$ -efficiency. Show that the same is true for  $A$ -efficiency and for  $G$ -efficiency.
- 9.15** Consider a  $k = 4$  scenario. The practitioner plans to use 20 design points in order to accommodate a complete second-order model. However, it is not clear what the true model is prior to conducting the experiment. As a result, the design selected should be one that is fairly **robust**—that is, one that will be reasonably efficient even if the model on which the design is based is not correct.
- Using the  $3^4$  factorial as the candidate list, construct the  $D$ -optimal design using a second-order model as the basis.
  - Suppose the edited model contains all linear terms and the interactions  $x_1x_2$  and  $x_1x_3$ . Give the  $D$ -efficiency of the design for this model.
  - Suppose, rather than use the computer-generated design, the practitioner uses a SCD containing
    - A PBD with 12 design points
    - Eight axial points with  $\alpha = 1.0$
 Give the  $D$ -efficiency of this design for a second-order model.
  - Give the  $D$ -efficiency of this design for the reduced model given in (b).
  - Comment on the above.
- 9.16** Consider Table 9.2. Notice that the CCD has high  $D$ -efficiencies but low  $G$ -efficiencies when the design has small numbers of center runs. Then when the number of center runs increases, the  $G$ -efficiency increases substantially. Explain why this is happening. Look at the variance dispersion graphs.
- 9.17** Consider a second-order model with four design variables and levels  $-2, -1, 0, 1, 2$  for each design variable. However, due to experimental constraints the candidate list should include:
- A  $3^4$  factorial with levels  $-1, 0, 1$
  - Eight axial points at  $\alpha = 2$
  - A center point
- Suppose 30 runs are to be used.
- Using a computer-generated design package, give the  $D$ -optimal design.
  - Compare your design in (a) with a CCD with  $\alpha = 2$  and  $n_c = 6$ . Use  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$  at the center of the design for your comparison.
  - Compare again, using the variances of the predicted values averaged over all of the candidate points.
  - Using VDG and FDS plots, compare where prediction is best and worst in each region, and determine which design is preferable.
- 9.18** Consider again the two designs compared in Exercise 9.17. Suppose that the “true” model involves all linear main effects and two-factor interactions only. Compare the  $D$ -efficiencies and the  $A$ -efficiencies.

- 9.19** Consider a hybrid 311B with three center runs. As an alternative, consider a BBD for  $k=3$  with  $n_0=3$  center runs. Answer the following questions:
- Which has the higher  $D$ -efficiency for a second-order model?
  - Which has the higher  $D$ -efficiency if, in fact, the true model is first-order First-order with two-factor interactions?
- 9.20** Show that for the case of a first-order model in  $k$  variables and spherical region of interest, a resolution III design with levels at  $\pm 1$  extremes is a  $I$ -optimal design. *Hint:* For this region,  $(1/K) \int_R x_i d\mathbf{x}$  and  $(1/K) \int_R x_i x_j d\mathbf{x}$  are both zero.
- 9.21** Consider the  $D$ -optimal design in Table 9.5 for the constrained region shown in Fig. 9.2. Suppose that you were to move all three replicates to the center of the design.
- Generate a graph of prediction standard errors similar to those in Fig. 9.18. What effect has moving the replicates to the center on those quantities?
  - Compute the  $D$ -efficiency of this new design relative to the one used in Example 9.13 (Table 9.5).
- 9.22** Reconsider the adhesive force experiment in Example 9.13. Find a  $D$ -optimal design with  $N=10$  runs. Use two runs as replicates. Compare the  $D$ -efficiency of this design with that of 12-run design in Table 9.5.
- 9.23** Consider a response surface design problem in  $k=3$  factors where the region is constrained so that  $x_1 + x_2 + x_3 \leq 2$ . Find a  $D$ -optimal design for fitting a second-order model with  $N=15$  runs, assuming that two runs are to be used as replicates. Generate plots of the standard deviation of the predicted response.
- 9.24** Rework Exercise 9.23 assuming that the constrained region is defined by  $-1.5 \leq x_1 + x_2 + x_3 \leq 2$ .
- 9.25** Consider the three experimental designs considered in Exercises 6.9, 6.10, and 6.11 for the auto-bumper plating process, assuming a full second-order model.
- Find the relative  $D$ -efficiencies of the designs in 6.10 and 6.11 relative to that in 6.9. Which design is best?
  - Find the  $G$ -efficiencies of the three designs.
  - Construct FDS plots for the three designs.
  - Based on the results of parts (a), (b), and (c), which design would you recommend?
- 9.26** Consider three competitors to the  $2^2$  factorial design, all operating on the same design space (using coded variables:  $x_1$  and  $x_2$  in  $[-1, 1]$ )

$$\mathbf{X}_A = \begin{bmatrix} 1 & 0.5 & 1 \\ 1 & -0.5 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{bmatrix} \quad \mathbf{X}_B = \begin{bmatrix} 1 & 0.5 & 0.5 \\ 1 & -0.5 & 0.5 \\ 1 & 1 & -0.5 \\ 1 & -1 & -0.5 \end{bmatrix} \quad \mathbf{X}_C = \begin{bmatrix} 1 & 0.5 & 0.5 \\ 1 & -0.5 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -0.5 \end{bmatrix}$$

For a first-order model,  $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$ , compare the three models and the  $2^2$  factorial design based on

- (a) orthogonality  
 (b) variance of the model coefficients  
 (c) scaled predicted variance (using some graphical summary)

Based on your findings in (a)–(c) rank them for each of the criteria.

- 9.27** Consider a situation in which there are two design variables and one is interested in fitting a second-order model. A curious and interesting comparison surfaces when one considers the central composite design (octagon) and a design with points placed at the vertices of a hexagon. Both are equiradial designs. The beauty of the hexagon (plus center runs)—or even the pentagon (plus center runs)—lies in its high relative efficiency in spite of its small design size. One should always choose the octagon unless it is too costly. However, it is interesting to compare the two *on a per observation basis*. Consider a comparison between a hexagon and a rotatable CCD. To make a valid comparison, one should scale the designs so they are comparable, that is, reside on the same radius (keep in mind that scaling is really arbitrary). The factorial points on the CCD are at radius  $\sqrt{2}$ . As scaled in Equation 9.7, the hexagon lies on a circle of radius 1. As a result, for proper comparison the hexagon should be multiplied by  $\sqrt{2}$ . Thus we have

$$\mathbf{D}_1 = \begin{matrix} & \begin{matrix} x_1 & x_2 \end{matrix} \\ \begin{bmatrix} \sqrt{2} & 0 \\ \sqrt{2}/2 & \sqrt{3/2} \\ -\sqrt{2}/2 & \sqrt{3/2} \\ -\sqrt{2} & 0 \\ -\sqrt{2}/2 & -\sqrt{3/2} \\ \sqrt{2}/2 & -\sqrt{3/2} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} & \text{(hexagon)} \end{matrix}$$

and, of course,

$$\mathbf{D}_1 = \begin{matrix} & \begin{matrix} x_1 & x_2 \end{matrix} \\ \begin{bmatrix} -1 & -1 \\ -1 & 1 \\ 1 & -1 \\ 1 & 1 \\ -\sqrt{2} & 0 \\ \sqrt{2} & 0 \\ 0 & -\sqrt{2} \\ 0 & \sqrt{2} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} & \text{(CCD)} \end{matrix}$$

Use  $\mathbf{D}_1$  and  $\mathbf{D}_2$  to construct scaled prediction variances. Both designs are rotatable, so a rather complete comparison can be made by filling in the following table:

| $\rho$     | $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ for $\mathbf{D}_1$ | $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ for $\mathbf{D}_2$ |
|------------|---|---|
| 0          |   |   |
| 0.5        |   |   |
| 1.0        |   |   |
| 1.5        |   |   |
| $\sqrt{3}$ |   |   |

Comment on your study.

- 9.28** Consider the comparison of the CCD and BBD shown in Fig. 9.5. Suppose that the total cost of the experiment is not critical to the decision of which design to select. Construct the VDG and FDS plots for the 17 run CCD and the 15 run BBD using the unscaled prediction variance  $\text{UPV} = \mathbf{x}^{(m)'}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}^{(m)}/\sigma^2$  to compare the designs. Which design would be preferred now?
- 9.29** Consider the hybrid designs 310 and 311B. By plotting FDS plots of both SPV and UPV for different numbers of center runs, determine what a suitable number of center runs would be for each design.
- 9.30** Consider the comparison of the  $3^3$  factorial design ( $N=27$ ) and the face-centered cube ( $N=15$ ) shown in Figs. 9.10 and 9.11. Construct the VDG and FDS plots for these designs using the unscaled prediction variance  $\text{UPV} = \mathbf{x}^{(m)'}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}^{(m)}/\sigma^2$  which does not penalize for design size to compare the designs. Which design would be preferred now? When would it be appropriate to use this comparison rather than the one based on the SPV?
- 9.31** Consider the comparison of the five-factor CCD on a cuboidal region using either the full factorial or the  $\frac{1}{2}$  fraction. Based on their relative  $D$ -,  $G$ -, and  $I$ -efficiencies, and the FDS plot for scaled prediction variance, which design is better?
- 9.32** Consider the hybrid designs 416A, 416B, and 416C, each with one center run. By plotting FDS plots of SPV, which design seems preferable?
- 9.33** Consider the Hoke  $\mathbf{D}_2$  ( $N=15$ ) and  $\mathbf{D}_6$  ( $N=19$ ) designs for four factors in a cuboidal region described in 8.1.1. Based on their relative  $D$ -,  $G$ -, and  $I$ -efficiencies, and the FDS plot for scaled prediction variance, which design is better?
- 9.34** For the same four factor cuboidal region described in Exercise 9.33, find  $I$ -optimal optimal designs of size  $N=15$  and  $N=19$  using a statistical software package. Based on their relative  $D$ -,  $G$ -, and  $I$ -efficiencies, and the FDS plot for scaled prediction variance, which design do you prefer? How do these designs compare to the Hoke  $\mathbf{D}_2$  and  $\mathbf{D}_6$  designs of Example 9.33?
- 9.35** An experiment was conducted to study the effect of temperature and type of oven on the life of a particular component being tested. Four types of ovens and three temperature levels were used in the experiment. Twenty-four pieces were assigned



**TABLE E9.1 The Experimental Design for Exercise 9.35**

| Temperature<br>(degrees) | Life                |                |                |                |
|--------------------------|---------------------|----------------|----------------|----------------|
|                          | Oven O <sub>1</sub> | O <sub>2</sub> | O <sub>3</sub> | O <sub>4</sub> |
| 500                      | 227                 | 214            | 225            | 260            |
|                          | 221                 | 259            | 236            | 229            |
| 550                      | 187                 | 181            | 232            | 246            |
|                          | 208                 | 179            | 198            | 273            |
| 600                      | 174                 | 198            | 178            | 206            |
|                          | 202                 | 194            | 213            | 219            |

randomly, two to each combination of treatments, and the results in Table E9.1 were recorded:

- (a) Fit an appropriate response surface model using  $-1, 0, 1$  coding for temperature. Be sure to determine if interaction exists among ovens and the response surface terms in temperature.
- (b) Find the estimated optimum temperature—that is, that which maximizes oven life. Does it depend on oven type?
- 9.36** To estimate the degree of a suspension polyethylene, we extract polyethylene by using a solvent, and we compare the gel contents (gel proportion). This method is called the **gel proportion estimation method**. In the book *Design of Experiments for Quality Improvement* published by the Japanese Standards Association (1989), a study is reported on the relationship between the amount of gel generated and three factors (solvent, extraction temperature, and extraction time) that influence amount of gel. The data are provided in Table E9.2.
- (a) Build (and state) an appropriate response surface model in terms of coded temperature and time. Write a separate model for each solvent if necessary.
- (b) Give an estimate of the conditions that maximize the amount of gel generated. Is there any indication of conditions for future experiments?
- 9.37** Consider a situation in which one is interested in three levels of each of two quantitative design variables. These levels are  $-1, 0, 1$ . In addition, there are three levels of a

**TABLE E9.2 The Experimental Design for Exercise 9.36**

| Solvent | Temperature | Amount of Gel |      |      |
|---------|-------------|---------------|------|------|
|         |             | Time: 4       | 8    | 16   |
| Ethanol | 120         | 94.0          | 93.8 | 91.1 |
|         |             | 94.0          | 94.2 | 90.5 |
|         | 80          | 95.3          | 94.9 | 92.5 |
|         |             | 95.1          | 95.3 | 92.4 |
| Toluene | 120         | 94.6          | 93.6 | 91.1 |
|         |             | 94.5          | 94.1 | 91.0 |
|         | 80          | 95.4          | 95.6 | 92.1 |
|         |             | 95.4          | 96.0 | 92.1 |

single qualitative factor. It is felt that the design should accommodate a second-order model in the quantitative variables and interaction between the linear terms and the qualitative variable.

- (a) Write out the model to be fitted.
- (b) Using 18 experimental runs, construct, using a computer-generated design package, an appropriate design.

**9.38** The following data was taken on two quantitative design variables and one two-level qualitative variable:

| $x_1$ | $x_2$ | $z$ | $y$  |
|-------|-------|-----|------|
| -1    | 1     | -1  | 17.5 |
| 1     | 1     | -1  | 22.1 |
| 0     | 0     | -1  | 25.2 |
| 0     | 0     | -1  | 26.5 |
| 0     | 0     | +1  | 29.4 |
| 0     | 0     | +1  | 31.3 |
| -1    | -1    | +1  | 19.2 |
| 1     | -1    | +1  | 13.4 |

- (a) Fit a model of the type

$$\hat{y} = b_0 + c_1z + b_1x_1 + b_2x_2 + b_{12}x_1x_2$$

- (b) Edit the model, and give recommendations on levels of  $x_1$  and  $x_2$  that maximize the response.

**9.39** In Section 6.3.3, the method of ridge analysis is used to optimize a process that is designed to convert 1,2-propanediol to 2,5-dimethyl-piperazine. The design is a central composite design with  $k=4$  and  $\alpha=1.4$ . Suppose, in addition, that two different catalysts are being used. The design and response values are in Table E9.3.

- (a) Construct an appropriate response surface model taking into account the two catalysts ( $z = +1, -1$ ).
- (b) Obtain a separate response surface model for each catalyst.

**9.40** Consider Exercise 3.9 in Chapter 3. Four factors were varied, and their effects were measured on alloy cracking. One of the factors is the type of heat treatment method (factor  $C$ ) at two levels. This is a good example of a mix of qualitative and quantitative variables. Use the coding  $-1$  and  $+1$  for all factors, including factor  $C$ .

- (a) Write out an appropriate model for relating the design variables to the response. Be sure the model is edited if necessary.
- (b) Use your result in (a) to develop two response surface models, one for each heat treatment method.

**9.41** Consider Exercise 4.12. A  $2^{6-3}$  fractional factorial is used to fit a first-order model

- (a) Write out the first-order model using  $-1, +1$  coding on the material type as the qualitative variable.
- (b) Write out separate models for material type 1 and material type 2.
- (c) What assumptions are you making in your developments in (b) above?

TABLE E9.3 The Experiment in Exercise 9.39

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $z$ | $y$  |
|-------|-------|-------|-------|-----|------|
| -1    | -1    | -1    | -1    | -1  | 68.4 |
| 1     | -1    | -1    | -1    | 1   | 28.5 |
| -1    | 1     | -1    | -1    | 1   | 22.0 |
| -1    | -1    | 1     | -1    | 1   | 15.7 |
| -1    | -1    | -1    | 1     | 1   | 48.5 |
| 1     | 1     | -1    | -1    | -1  | 28.5 |
| 1     | -1    | 1     | -1    | -1  | 18.7 |
| 1     | -1    | -1    | 1     | -1  | 65.8 |
| -1    | 1     | 1     | -1    | -1  | 9.2  |
| -1    | 1     | -1    | 1     | -1  | 32.9 |
| -1    | -1    | 1     | 1     | -1  | 12.5 |
| -1    | 1     | 1     | 1     | 1   | 49.8 |
| 1     | -1    | 1     | 1     | 1   | 12.9 |
| 1     | 1     | -1    | 1     | 1   | 30.9 |
| 1     | 1     | 1     | -1    | 1   | 22.9 |
| 1     | 1     | 1     | 1     | -1  | 28.2 |
| -1.4  | 0     | 0     | 0     | -1  | 32.1 |
| 1.4   | 0     | 0     | 0     | -1  | 18.2 |
| 0     | -1.4  | 0     | 0     | -1  | 19.2 |
| 0     | 1.4   | 0     | 0     | -1  | 50.5 |
| 0     | 0     | -1.4  | 0     | 1   | 52.1 |
| 0     | 0     | 1.4   | 0     | 1   | 39.2 |
| 0     | 0     | 0     | -1.4  | 1   | 30.7 |
| 0     | 0     | 0     | 1.4   | 1   | 40.7 |
| 0     | 0     | 0     | 0     | 1   | 39.5 |
| 0     | 0     | 0     | 0     | 1   | 38.2 |

- 9.42** Consider the data in Exercise 4.14. Two qualitative variables appear.
- Fit a model containing all main effects and two-factor interactions.
  - Edit the model to contain only significant terms.
  - Use your model in (b) to write separate models for each combination of the qualitative variables.
- 9.43** There are 27 nonisomorphic fractional factorial designs for 16-run experiments involving six factors. The appendix of Johnson and Jones (2010) lists the defining equations of the different choices. The results from Lu, Johnson, and Anderson-Cook (2014) evaluating two design criteria,  $E(s^2)$  and  $\text{tr}(\mathbf{AA}')$ , for all of the possible designs are shown in Table E9.4. Recall  $E(s^2)$  looks at the correlation between all of the main and two-factor interactions among the six factors and  $\text{tr}(\mathbf{AA}')$  considers the amount of bias on the main effects, if some of the two-factor interactions are active.
- Which is the best design for minimizing  $E(s^2)$ ? Why might an experimenter select this design?
  - Which design is best for minimizing  $\text{tr}(\mathbf{AA}')$ ? Why might an experimenter prefer this design?

**TABLE E9.4 Comparison of Six-Factor 16-Run Experiments**

| Design # | $E(s^2)$ | $\text{tr}(\mathbf{AA}')$ | Design # | $E(s^2)$ | $\text{tr}(\mathbf{AA}')$ |
|----------|----------|---------------------------|----------|----------|---------------------------|
| 1        | 25.6     | 12                        | 15       | 9.14     | 4.5                       |
| 2        | 10.97    | 6                         | 16       | 18.29    | 6                         |
| 3        | 7.31     | 6                         | 17       | 10.97    | 6                         |
| 4        | 7.31     | 3                         | 18       | 7.31     | 6                         |
| 5        | 10.97    | 0                         | 19       | 9.14     | 3                         |
| 6        | 7.31     | 3.75                      | 20       | 9.14     | 4.5                       |
| 7        | 9.14     | 4.5                       | 21       | 14.63    | 7.5                       |
| 8        | 9.14     | 1.5                       | 22       | 10.97    | 4.5                       |
| 9        | 9.14     | 5.25                      | 23       | 9.14     | 5.25                      |
| 10       | 18.29    | 9                         | 24       | 9.14     | 4.5                       |
| 11       | 10.97    | 6                         | 25       | 14.63    | 6                         |
| 12       | 10.97    | 3                         | 26       | 10.97    | 6                         |
| 13       | 7.31     | 3                         | 27       | 12.8     | 6                         |
| 14       | 7.31     | 3                         |          |          |                           |

- (c) Examine a scatterplot of the two design criteria and use it to construct the Pareto front described in Chapter 7 with the set of best designs based on  $E(s^2)$  and  $\text{tr}(\mathbf{AA}')$ ?
- (d) If an experimenter wanted to have balance between the two criteria, which design would you recommend?
- (e) If an experimenter wanted to select design 10 to run, how would you convince him that this was not a sensible choice based on these two criteria? Which design might be a better alternative?
- 9.44** Consider Exercise 9.43. Designs 1, 2, 3, 4, 5, 10, 16, 21, 25 are regular designs (standard fractional factorials, or replicated fractional factorials) and the remainder of the designs are nonregular (see Section 4.8). Looking at separate scatterplots for the two classes of designs,
- (a) Identify the Pareto front based on  $E(s^2)$  and  $\text{tr}(\mathbf{AA}')$  for the regular designs.
- (b) Identify the Pareto front based on  $E(s^2)$  and  $\text{tr}(\mathbf{AA}')$  for the nonregular designs.
- (c) If an experimenter wanted to choose a best design from between regular and nonregular designs, which class of designs would you recommend? Explain.
- (d) Based on your answer in (c), which design would you select to run if you had an experiment with 16 runs and six factors?
- 9.45** Consider Exercise 9.43. Using the constrained optimization described in Section 7.2.3 to select best designs
- (a) with the best  $E(s^2)$ , conditional on  $\text{tr}(\mathbf{AA}')$  being no worse than 2.5.
- (b) with the best  $\text{tr}(\mathbf{AA}')$ , conditional on  $E(s^2)$  being no worse than 8.
- (c) with the best  $\text{tr}(\mathbf{AA}')$ , conditional on  $E(s^2)$  being within 30% of the best possible value.

TABLE E9.5 Comparison of Seven-Factor 16-Run Experiments

| Design # | Regular | $E(s^2)$ | $\text{tr}(\mathbf{AA}')$ | Design # | Regular | $E(s^2)$ | $\text{tr}(\mathbf{AA}')$ |
|----------|---------|----------|---------------------------|----------|---------|----------|---------------------------|
| 1        | Y       | 28.44    | 21                        | 29       | N       | 10.67    | 8.25                      |
| 2        | Y       | 14.22    | 12                        | 30       | N       | 10.16    | 7.5                       |
| 3        | Y       | 12.19    | 9                         | 31       | N       | 12.19    | 6                         |
| 4        | Y       | 10.16    | 9                         | 32       | N       | 10.16    | 6                         |
| 5        | Y       | 10.16    | 6                         | 33       | N       | 10.16    | 6                         |
| 6        | Y       | 14.22    | 0                         | 34       | N       | 11.17    | 9                         |
| 7        | N       | 10.16    | 7.5                       | 35       | N       | 10.67    | 8.25                      |
| 8        | N       | 10.67    | 8.25                      | 36       | N       | 10.67    | 6.75                      |
| 9        | N       | 12.19    | 9                         | 37       | N       | 10.16    | 7.5                       |
| 10       | N       | 10.16    | 9                         | 38       | N       | 12.19    | 9                         |
| 11       | N       | 10.16    | 6                         | 39       | N       | 10.16    | 9                         |
| 12       | N       | 12.19    | 3                         | 40       | N       | 12.19    | 10.5                      |
| 13       | N       | 10.16    | 9                         | 41       | Y       | 16.25    | 12                        |
| 14       | N       | 12.19    | 10.5                      | 42       | Y       | 14.22    | 12                        |
| 15       | N       | 10.16    | 7.5                       | 43       | N       | 10.16    | 7.5                       |
| 16       | Y       | 20.32    | 15                        | 44       | N       | 14.22    | 9                         |
| 17       | N       | 14.22    | 12                        | 45       | N       | 12.19    | 9                         |
| 18       | N       | 12.19    | 9                         | 46       | N       | 10.16    | 7.5                       |
| 19       | N       | 10.16    | 9                         | 47       | N       | 10.16    | 9                         |
| 20       | N       | 14.22    | 6                         | 48       | N       | 11.17    | 9                         |
| 21       | N       | 10.16    | 6                         | 49       | N       | 11.17    | 6                         |
| 22       | N       | 10.16    | 6                         | 50       | N       | 11.17    | 9                         |
| 23       | N       | 10.67    | 6.75                      | 51       | Y       | 14.22    | 10.5                      |
| 24       | N       | 12.19    | 9                         | 52       | Y       | 14.22    | 10.5                      |
| 25       | N       | 10.16    | 9                         | 53       | N       | 10.16    | 7.5                       |
| 26       | N       | 11.17    | 4.5                       | 54       | N       | 10.67    | 8.25                      |
| 27       | N       | 11.17    | 7.5                       | 55       | N       | 10.67    | 6.75                      |
| 28       | N       | 11.17    | 4.5                       |          |         |          |                           |

**9.46** There are 55 nonisomorphic fractional factorial designs for 16-run experiments involving seven factors. Lu, Johnson, and Anderson-Cook (2014) evaluated two design criteria,  $E(s^2)$  and  $\text{tr}(\mathbf{AA}')$ , and the results for all of the possible designs are shown in Table E9.5.

- Using a scatterplot of the two criteria  $E(s^2)$  and  $\text{tr}(\mathbf{AA}')$ , is the best design to minimize  $E(s^2)$  regular or nonregular? What about for  $\text{tr}(\mathbf{AA}')$ ?
- Construct the Pareto front based on all of the designs based on  $(s^2)$  and  $\text{tr}(\mathbf{AA}')$ .
- Are there more regular or nonregular designs on the Pareto front?
- What are some of the worst performing designs based on these criteria? How much worse are they than choices on the Pareto front? Describe how this would impact the results of the experiment if one of the suboptimal designs were selected.
- There are several ties on the Pareto front (different designs with the same  $E(s^2)$  and  $\text{tr}(\mathbf{AA}')$  values). How would you suggest that an experimenter choose between the alternatives?

- 9.47** Consider Exercise 9.46. Using the Pareto front obtained in 9.45(b), use the desirability function approach of Section 7.2.3 to select a best design.
- If the experimenter scales the individual desirability functions so that the best value on the Pareto front has value 1 and the worst value on the front has value 0, which design would be identified as best if the experimenter wanted to use an additive desirability function to combine the two criteria with equal weight the contributions?
  - Repeat (a) with the individual desirability scales chosen so the best and worst values in Table E9.5 define 1 and 0, respectively.
  - Suppose that the experimenter wanted to prioritize  $E(s^2)$  and give it twice the weight of  $\text{tr}(\mathbf{AA}')$ . What would be the best design using the same scaling as described in (a)? Using the scaling in (b)?
  - Describe which of the scaling seems more appealing and how you would recommend the experimenter choosing a best scale.
- 9.48** Consider the two Definitive Screening Designs ( $m = 4$ ,  $c = 2$ ) for 4 continuous variables ( $A, B, C, D$ ) and two categorical (or qualitative) factors ( $E, F$ ) in Section 9.3.5. The DSD-augment design has fewer runs than then ORTH-augment design.
- Calculate the determinant of the moment matrix,  $|\mathbf{M}| = |\mathbf{X}'\mathbf{X}|/N^p$  for each design assuming that a first order model is fit for all six factors. Which design has a higher  $D$ -efficiency? What is the relative efficiency of the poorer design,  $|\mathbf{M}_1|/|\mathbf{M}_2|$ ?
  - The criterion  $|\mathbf{M}|$  penalizes larger designs. If this penalty is removed, by looking at only the numerator, what is the relative efficiency of the poorer design now?
  - Suppose that after model fitting is performed, the final model includes the terms,  $A, B, C, E, F, AB, AE, BF, CE, EF$ . Calculate  $|\mathbf{M}| = |\mathbf{X}'\mathbf{X}|/N^p$  for this model.
  - By examining  $\mathbf{X}'\mathbf{X}$ , can all of the terms in the model in (a) be estimated independently for each design? What is the consequence of this result?
  - By examining  $\mathbf{X}'\mathbf{X}$ , can all of the terms in the model in (b) be estimated independently for each design? What is the consequence of this result?

### Short Answer Questions

- 9.49** Answer the following short answer questions:
- What are some of the trade-offs between using a central composite design and a design with a smaller run size?
  - Justify why a  $D$ -optimal design is a good choice for designed experiment? What is the primary attribute of the design that this criterion focuses on?
  - What are the trade-offs between focusing on  $G$ - and  $I$ -optimality for good prediction variance properties of a design?
  - Describe the impact of changing the number of center runs in a central composite or Box–Behnken design.
  - Describe the difference between the variance dispersion graph and the fraction of design space plot for summarizing prediction variance.

- (f) Describe a situation where using a computer-generated design might lead to a better choice than using a standard choice of design.
- (g) To use an algorithm to create a computer-generated design, what choices does the experimenter need to specify?
- (h) Describe the difference between point exchange and coordinate exchange search algorithms.
- (i) Describe why different strategies and criteria are needed when qualitative factors are considered as part of a designed experiment.
- (j) Justify why using a Pareto front approach to design optimization might be advantageous in many situations.

**9.50** Answer the following True/False questions:

- (a) If the  $\text{Max Var}[\hat{y}(\mathbf{x})] = 30$  for a second-order model in five variables, the  $G$ -efficiency of the experimental design is 0.75 or 75%.
- (b) Any design that is  $D$ -optimal is also  $G$ -optimal.
- (c) A variance dispersion graph is a two-dimensional representation of the prediction variance properties of an experimental design when that design is used to fit a specific model.
- (d) Small composite designs for fitting response surfaces can be made rotatable by proper choice of  $\alpha$ .
- (e) Small composite designs generally have inferior prediction variance properties in comparison to hybrid designs.
- (f) Fraction of design space plots can be constructed for any shaped design region.
- (g) Variance dispersion graphs can be adapted for any shaped design region.
- (h) A design is  $G$ -optimal if the maximum value of the scaled prediction variance over the design region is exactly equal to the number of runs in the design,  $n$ .
- (i) Point exchange and coordinate exchange are two methods for finding  $D$ -optimal designs.
- (j)  $D$ -optimal designs are usually found by exchange algorithms, which are guaranteed to produce a globally-optimal design.
- (k) A design is used to fit a second-order model in five variables. The maximum scaled prediction variance over the design region for the model is 24. The  $G$ -efficiency for this design is 85%.
- (l)  $I$ -optimality (or  $Q$  or  $IV$  optimality) is a better criterion than  $D$ -optimality for selecting a second-order response surface design.
- (m) Consider two competing 20-run designs. For design 1, we have found that  $|(\mathbf{X}'\mathbf{X})^{-1}| = 5\text{E-}9$ , and for design 2 we have found that  $|(\mathbf{X}'\mathbf{X})^{-1}| = 4\text{E-}9$ . Both designs are for  $k = 4$  variables and will be used to fit a model with 15 parameters. The relative efficiency of design 1 to design 2 is 80%.
- (n) A design is  $A$ -optimal if the maximum value of the scaled prediction variance over the design region is exactly equal to the number of model parameters,  $p$ .
- (o) A primary disadvantage of hybrid designs for fitting response surfaces is that they require unusual levels for the design factors.
- (p) Adding one or two center runs to a hybrid design is a good idea because it stabilizes the prediction variance over the design region.

- (q) Designs which are perfectly spherical (all design points equidistant from the design center), with no center runs, do not give good prediction variance at the center of the region.
- (r) The “messy levels” associated with hybrid designs are a real problem as even minor deviations from specified levels results in poor design efficiency.

**9.51** Circle the correct answer for each question:

- (a) The variance dispersion graph for a design
  - (i) Always only has one line to represent the prediction variance.
  - (ii) Can only be used with second-order rotatable designs.
  - (iii) Deals with the scaled prediction variance on a sphere of radius  $r$  but ignore the volume associated with this information.
  - (iv) None of (i)–(iii) are correct statements.
  - (v) All of (i)–(iii) are correct statements.
- (b) The fraction of design space plot for a design
  - (i) Always only has one line to represent the prediction variance.
  - (ii) Can only be used with second order rotatable designs.
  - (iii) Can only be constructed for cuboidal design regions.
  - (iv) None of (i)–(iii) are correct statements.
  - (v) All of (i)–(iii) are correct statements.



---

# 10

---

## ADVANCED TOPICS IN RESPONSE SURFACE METHODOLOGY

In this chapter we deal with several more advanced response surface topics. This chapter focuses on the effects of bias in the fitting of response surface models, the effects of errors in the design variables, experiments with computer models, minimum-bias estimation as an alternative to the method of least squares, the role of artificial neural networks in RSM, response surface designs in a split-plot structure, and how to deal with non-normal responses through the use of generalized linear models.

### 10.1 EFFECTS OF MODEL BIAS ON THE FITTED MODEL AND DESIGN

In Chapter 8 we discussed the role of model bias on the test of lack of fit and the estimate of the error variance  $\sigma^2$ . We learned that if the fitted polynomial approximation is given by

$$\hat{y} = \mathbf{X}_1 \mathbf{b}_1 \quad (10.1)$$

and the true model for the expected response  $E(\mathbf{y})$  is better approximated by

$$E(\mathbf{y}) = \mathbf{X}_1 \beta_1 + \mathbf{X}_2 \beta_2 \quad (10.2)$$

then the estimator  $\mathbf{b}_1$  is biased, and the estimator of  $\sigma^2$ , the error mean square in the analysis of variance, is biased upward. In fact, it is the bias in the mean squared error that one is detecting in a standard lack-of-fit test. This is the result of the mean squared error being a combination of bias introduced from a misspecified model and pure error, if the design contains replicates.

---

*Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, Fourth Edition, by Raymond H. Myers, Douglas C. Montgomery, and Christine M. Anderson-Cook. Copyright © 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

The important point here is that there are two types of model errors. In the postulated model

$$\mathbf{y} = \mathbf{X}_1 \boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}$$

$\mathbf{X}_1 \boldsymbol{\beta}_1$  is merely an approximation; errors associated with the model can be put into two categories:

1. Random errors—that is, elements in  $\boldsymbol{\varepsilon}$  (with variance  $\sigma^2$ )
2. Systematic errors—that is, bias errors of the type  $\mathbf{X}_1 \boldsymbol{\beta}_1 - E(\mathbf{y})$

Apart from the simple lack-of-fit test, the entire emphasis in design and analysis has been placed on variability produced by random variation. In particular, all of experiment design assessment criteria in Chapter 9 deal with variance oriented criteria; that is,  $D$ ,  $G$ ,  $A$ ,  $V$ , and  $I$  (or  $IV$ ) optimality all deal with criteria that relate to  $\mathbf{X}'\mathbf{X}$  or the moment matrix  $\mathbf{X}'\mathbf{X}/N$  for a particular model that is assumed to be correct. In what follows we link the  $I$ -optimality criterion with another involving model bias.

**What Effect Does Design Have on Variance and Bias?** The most effective way to demonstrate the effect of design on variance and bias is to illustrate with a simple scenario. Suppose that  $k = 1$  and the **true model**  $f(x)$  is given by

$$f(x) = \beta_0 + \beta_1 x + \beta_{11} x^2 = 1 + x + x^2 \quad (10.3)$$

and it is approximated by a straight line with the formal fitted model

$$\hat{y}(x) = b_0 + b_1 x$$

For simplicity we assume that the region of interest and region of operability are the same, namely, the interval  $[-1, +1]$ . In addition, suppose the design contains a single run at each of the endpoints,  $-1$  and  $+1$ . As a result, the variance transmitted to the predicted value, namely  $v(x) = N \text{Var}[\hat{y}(x)]$ , is given by (assuming  $\sigma = 1$ )

$$\begin{aligned} v(x) &= 2[\text{Var}(b_0) + x^2 \text{Var}(b_1)] \\ &= 2 \left[ \frac{1}{2} + x^2 \text{Var}(b_1) \right] \\ &= 2 \left[ \frac{1}{2} + \frac{x^2}{2} \right] \end{aligned} \quad (10.4)$$

because  $\text{Var}(b_1) = \sigma^2 / \sum_{i=1}^2 (x_i)^2 = \frac{1}{2}$ . As a result,

$$v(x) = 1 + x^2 \quad (10.5)$$

The measure of bias that is consistent in units with  $v(x)$  is given by

$$d(x) = N[E\hat{y}(x) - f(x)]^2 \quad (10.6)$$

We have  $E[\hat{y}(x)] = E(b_0) + xE(b_1)$ . Recall from Chapter 8 that if one fits an incomplete model  $\hat{y} = \mathbf{X}_1\boldsymbol{\beta}_1$ , when in fact  $E(\mathbf{y}) = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2$ , then

$$E(\mathbf{b}_1) = \boldsymbol{\beta}_1 + (\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1\mathbf{X}_2\boldsymbol{\beta}_2$$

Indeed, in our case

$$\boldsymbol{\beta} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad (\mathbf{X}'_1\mathbf{X}_1)^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}, \quad \text{and}$$

$$\mathbf{X}'_1\mathbf{X}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} (1)^2 \\ (-1)^2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

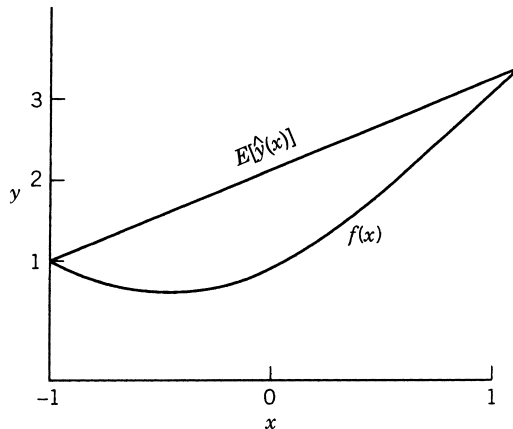
Thus, because  $\boldsymbol{\beta}_2 = \beta_{11} = 1$ , we obtain

$$E \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} \beta_0 + 1 \\ \beta_1 \end{bmatrix}$$

As a result, in this case the bias is transmitted only to the intercept estimate. Thus  $E[\hat{y}(x)] = \beta_0 + 1 + \beta_1x = 2 + x$ . From Equation 10.6 we obtain

$$\begin{aligned} d(x) &= 2[\beta_0 + 1 + \beta_1x - (\beta_0 + \beta_1x + x^2)]^2 \\ &= 2[1 - x^2]^2 \end{aligned} \tag{10.7}$$

The results in Equations 10.5 and 10.7 represent the two components that contribute to error in predicted values  $\hat{y}(x)$  in the region  $[-1, +1]$ . Now, consider Fig. 10.1. The actual function  $f(x)$  is plotted along with  $E[\hat{y}(x)] = 2 + x$ . The differences, or vertical deviations, represent **model bias** values that are incurred with this endpoint design, namely, the two-point design at  $-1$  and  $+1$ .



**Figure 10.1** True function  $f(x)$  and  $E[\hat{y}(x)]$  for design with points at  $-1, +1$ .

Now, suppose we consider an alternative design, namely, a two-point design at the points  $-0.6$  and  $+0.6$ . Again, in the fitted model  $\hat{y}(x) = b_0 + b_1x$ ,  $b_1$  is found to be unbiased for  $\beta_1$ , but  $b_0$  is biased. Following the same steps as before, we have

$$E \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{0.5}{(0.6)^2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -0.6 & 0.6 \end{bmatrix} \begin{bmatrix} (-0.6)^2 \\ (0.6)^2 \end{bmatrix} \beta_{11} = \begin{bmatrix} \beta_0 + 0.36 \\ \beta_1 \end{bmatrix}$$

Note that the bias in the intercept is not as great as in the case of the previous endpoint design. Thus the squared bias  $d(x)$  is given by

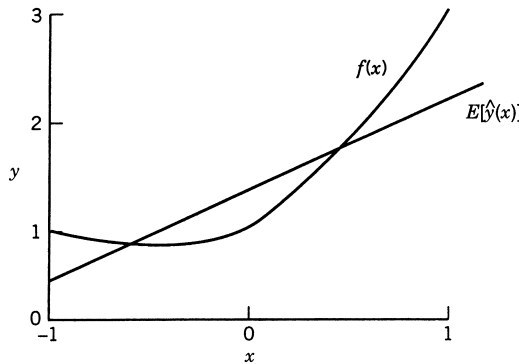
$$\begin{aligned} d(x) &= 2[E(b_0 + b_1x) - (\beta_0 + \beta_1x + \beta_{11}x^2)]^2 \\ &= 2[\beta_0 + 0.36 + \beta_1x - \beta_0 - \beta_1x - \beta_{11}x^2]^2 \\ &= 2[0.36 - x^2]^2 \end{aligned}$$

The expected value of the fitted approximation, namely  $E[\hat{y}(x)]$ , is given by

$$E[\hat{y}(x)] = 1.36 + x$$

Figure 10.2 displays  $f(x)$  and  $E[\hat{y}(x)]$  in the case of the second two-point design.

Figures 10.1 and 10.2 reveal how the structure of the bias values  $E[\hat{y}(x)] - f(x)$  are affected by the two designs in different ways. In the case discussed here the endpoint design results in rather small bias errors only near the endpoints where the design points are located. Near the design center, very large bias errors occur because there are no data there. As a result, it would appear reasonable that if design points are taken closer to the center, as in the case of the second design, the bias errors should be smaller. In the case of the second design, very small errors occur near the design points and moderate errors are experienced near the design center and at the design perimeter. Clearly, designs with points that are pushed closer to the center of the design afford considerably more protection against bias errors than does the endpoint design, though of course the latter is a minimum variance design. This suggests that when one fears model inadequacy, perhaps there is a



**Figure 10.2** True function  $f(x)$  and  $E[\hat{y}(x)]$  for design at points  $-0.6, 0.6$ .

need to consider some type of compromise design, because a strategy involving protection against bias errors works counter to a strategy designed to minimize variance.

This trade-off between the variance of a design and potential bias is often referred to as the *bias-variance trade-off*. Recall in Section 8.1 we discussed 11 desirable properties of response surface designs. We divided the properties into different categories, where the first was based on the assumption that the experimenter had made an appropriate choice of model on which to base the choice of design. If the model is correct (or closely approximates the true underlying relationship between the response and factors), then focusing on good estimation and prediction based on that model is the right priority.

However, if the model is not correct (or is a poor approximation to the true underlying relationship), then the second category described in Section 10.1 becomes much more important. In this case, the focus is on minimizing the consequences of the misspecified model, by protecting against bias and having the opportunity to discover that the assumed model does not perform well.

Where this becomes important for the experimenter is that decisions about the assumed model must be made *before* any data are collected. Hence, this highlights the need to think carefully about the choice of model, as much depends on this choice throughout the analysis of data obtained. We should note that both the designs shown in Figs. 10.1 and 10.2 do not have any opportunity to detect the curvature of the true underlying relationship defined in Equation 10.3. A third possible design might be to collect data at three locations  $-1, 0, +1$  or  $-0.6, 0, +0.6$ . Either of these designs would now provide information about the curvature and could provide the opportunity to update the model. However, including this additional design location does require an additional observation to be collected. Hence, in addition to thinking about the bias–variance trade-off, some consideration of cost now has also been added.

It should be increasingly clear to the reader that selecting a good design has many facets to balance and that naively assuming that the chosen model is correct can have negative consequences, if that assumption later proves to be incorrect. An approach such as that described in Section 9.4 might be considered to formally examine the bias–variance trade-off if there is model uncertainty.

## 10.2 A DESIGN CRITERION INVOLVING BIAS AND VARIANCE

The scientist or engineer who applies RSM may not always be aware when he or she should fit a second-order model. However, very often only a moderate amount of model inadequacy (in magnitude less than is detectable by a lack-of-fit test) can lead to errors that are as sizeable as variation due to random errors.

It is natural to consider combining bias and variance through the **mean squared error** criterion

$$E[\hat{y}(\mathbf{x}) - f(\mathbf{x})]^2 \quad (10.8)$$

Here  $\mathbf{x}$  is a vector to allow for multiple design variables. Box and Draper (1959, 1963) introduced an average mean squared error criterion in which the mean squared error in Equation 10.8 is averaged over a region of interest in the design variables. They also showed some compelling evidence that bias due to model misspecification should not be ignored when one chooses an experimental design for response surface exploration. Indeed,

there are situations when model bias is the dominant component in the mean squared error formulation. As we saw in the development in Section 10.1, it often occurs that the minimum variance approach applies a premium to pushing points to the edge of the region of interest (much as in a  $D$ -optimal approach), whereas a minimum bias approach calls for more placement of points that are at an intermediate distance from the design center. The details depend a great deal on (a) the model that one is fitting and (b) the model that one chooses to protect against. The latter model usually involves model terms that are of order higher than that of the assumed model.

**Generalization of the Average Mean Squared Error** Consider the mean squared error of Equation 10.8. This **average squared error loss criterion** is written so as to be a function of  $\mathbf{x}$ , the location at which one is predicting the response with  $\hat{y}(\mathbf{x})$ . Now, the mean squared error in Equation 10.8 can be partitioned as

$$\begin{aligned} E[\hat{y}(\mathbf{x}) - f(\mathbf{x})]^2 &= E\{[\hat{y}(\mathbf{x}) - E[\hat{y}(\mathbf{x})]] + [E[\hat{y}(\mathbf{x})] - f(\mathbf{x})]\}^2 \\ &= E\{\hat{y}(\mathbf{x}) - E[\hat{y}(\mathbf{x})]\} + \{E[\hat{y}(\mathbf{x})] - f(\mathbf{x})\}^2 \end{aligned} \quad (10.9)$$

The cross-product term in the above squaring operation is zero. The two terms on the right-hand side of Equation 10.9 are, respectively, the variance and the squared bias of  $\hat{y}(\mathbf{x})$ .

It should be clear from the foregoing and other developments in this text that there is a tradeoff in dealing with variance and bias. It should also be apparent that Equation 10.9 does not afford the user a single-number criterion, because the mean squared error depends on  $\mathbf{x}$ , the point at which one predicts. However, Box and Draper suggested the use of an **average mean squared error** (AMSE), where the quantity  $E[\hat{y}(\mathbf{x}) - f(\mathbf{x})]^2$  in Equation 10.9 is averaged over  $\mathbf{x}$  in a preselected region of interest. In other words, a quantity

$$\text{AMSE} = \frac{NK}{\sigma^2} \int_R E[\hat{y}(\mathbf{x}) - f(\mathbf{x})]^2 d\mathbf{x} \quad (10.10)$$

was offered as a single-number criterion. The integral  $\int_R$ , along with the multiplication by  $K = 1/\int_R d\mathbf{x}$ , the reciprocal of the volume of the region, implies that the quantity AMSE is an average of the MSE over  $R$ . The quantity  $N/\sigma^2$  is a scale factor in the same spirit as that provided by the scaled prediction variance  $N \text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$ , which has been used throughout the text. The AMSE can be subdivided as follows:

$$\begin{aligned} \text{AMSE} &= \frac{NK}{\sigma^2} \int_R E\{\hat{y}(\mathbf{x}) - E[\hat{y}(\mathbf{x})]\}^2 d\mathbf{x} + \frac{NK}{\sigma^2} \int_R \{E[\hat{y}(\mathbf{x})] - f(\mathbf{x})\}^2 d\mathbf{x} \\ &= \text{APV} + \text{ASB} \end{aligned} \quad (10.11)$$

where APV stands for **average prediction variance** and ASB stands for **average squared bias**. Here, of course, the averaging is done over the region of interest  $R$ . In what follows we focus heavily on designs that minimize APV and ASB, with strong emphasis on the latter. Note that minimization of APV is equivalent to the  $I$ -optimality criterion discussed in Chapter 9.

The partition of mean squared error into variance and squared bias should bring to mind the material in Chapter 8 dealing with lack of fit. The structure of bias stems from the assumption that the model

$$\hat{\mathbf{y}} = \mathbf{X}_1 \mathbf{b}_1$$

is fitted and the true model is a polynomial of a higher degree. The matrix  $\mathbf{X}_1$  contains model terms of a low-order approximation of the true structure, and the matrix  $\mathbf{X}_2$  contains additional higher-order terms that one wishes to protect against in designing the experiment. Thus

$$\mathbf{f} = E(\mathbf{y}) = \mathbf{X}_1 \boldsymbol{\beta}_1 + \mathbf{X}_2 \boldsymbol{\beta}_2$$

Here, of course,  $\mathbf{y}$  is the  $N \times 1$  vector of observations. From material in Chapter 8, we know that the estimator  $\mathbf{b}_1$  is biased, namely,

$$E(\mathbf{b}_1) = \boldsymbol{\beta}_1 + \mathbf{A} \boldsymbol{\beta}_2 \quad (10.12)$$

where

$$\mathbf{A} = (\mathbf{X}'_1 \mathbf{X}_1)^{-1} \mathbf{X}'_1 \mathbf{X}_2$$

As a result, the bias portion of AMSE, namely ASB, can be written

$$\text{ASB} = \frac{NK}{\sigma^2} \int_R \{ \mathbf{x}'_1 E(\mathbf{b}_1) - [\mathbf{x}'_1 \boldsymbol{\beta}_1 + \mathbf{x}'_2 \boldsymbol{\beta}_2] \}^2 d\mathbf{x} \quad (10.13)$$

where  $\mathbf{x}_1$  contains the  $p_1$  model terms that are fitted, and  $\mathbf{x}_2$  contains the  $p_2$  terms that are ignored. On invoking Equation 10.12, Equation 10.13 reduces to

$$\begin{aligned} \text{ASB} &= \frac{NK}{\sigma^2} \int_R [(\mathbf{x}'_1 \mathbf{A} - \mathbf{x}'_2) \boldsymbol{\beta}_2]^2 d\mathbf{x} \\ &= \frac{NK}{\sigma^2} \int_R \boldsymbol{\beta}'_2 [\mathbf{A}' \mathbf{x}_1 - \mathbf{x}_2] [\mathbf{x}'_1 \mathbf{A} - \mathbf{x}'_2] \boldsymbol{\beta}_2 \\ &= \frac{NK}{\sigma^2} \int_R [\boldsymbol{\beta}'_2 \mathbf{A}' \mathbf{x}_1 \mathbf{x}'_1 \mathbf{A} \boldsymbol{\beta}_2 - 2 \boldsymbol{\beta}'_2 \mathbf{x}_2 \mathbf{x}'_1 \mathbf{A} \boldsymbol{\beta}_2 + \boldsymbol{\beta}'_2 \mathbf{x}_2 \mathbf{x}'_2 \boldsymbol{\beta}_2] d\mathbf{x} \end{aligned} \quad (10.14)$$

For the case of APV we have

$$\text{APV} = \frac{NK}{\sigma^2} \int_R \mathbf{x}'_1 (\mathbf{X}'_1 \mathbf{X}_1)^{-1} \mathbf{x}_1 d\mathbf{x}$$

**10.2.1    The Case of a First-Order Fitted Model and Cuboidal Region**

For the situation where the fitted model is first-order, we have

$$\hat{y}(\mathbf{x}) = b_0 + \sum_{i=1}^k b_i x_i$$

and hence  $\mathbf{x}'_1 = [1, x_1, \dots, x_k]$ . If we further assume that we are interested in protecting against a true model that is of second order, then

$$\mathbf{X}_1 = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{N,1} & \cdots & x_{N,k} \end{bmatrix}$$

and

$$\mathbf{X}_2 = \begin{bmatrix} x_{11}^2 & \cdots & x_{1k}^2 & x_{11}x_{12} & \cdots & x_{1,k-1} & x_{1,k} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ x_{N,1}^2 & \cdots & x_{N,k}^2 & x_{N,1}x_{N,2} & \cdots & x_{N,k-1} & x_{N,k} \end{bmatrix} \tag{10.15}$$

Assume that both the region of interest and the region of operability on the design variables is  $-1 \leq x_i \leq 1$  for  $i = 1, 2, \dots, k$ . If we consider variance alone, we can borrow from Chapter 9, where we learned that the *I*-optimal design is an orthogonal design in which all levels are set at the  $\pm 1$  extremes. However, for the case of bias, one would expect from the discussion in Section 10.1 that levels should be placed inward from the  $\pm 1$  extremes.

Suppose we consider the minimization of ASB in Equation 10.14 by choice of design. Using the fact that  $\mathbf{A} = (\mathbf{X}'_1 \mathbf{X}_1)^{-1} (\mathbf{X}'_1 \mathbf{X}_2) = \mathbf{M}_{11}^{-1} \mathbf{M}_{12}$ , where  $\mathbf{M}_{11} = \mathbf{X}'_1 \mathbf{X}_1 / N$  and  $\mathbf{M}_{12} = \mathbf{X}'_1 \mathbf{X}_2 / N$  are **moment matrices**. We have

$$ASB = \frac{NK}{\sigma^2} \int_R \boldsymbol{\beta}'_2 [\mathbf{M}'_{12} \mathbf{M}_{11}^{-1} \mathbf{x}_1 \mathbf{x}'_1 \mathbf{M}_{11}^{-1} \mathbf{M}_{12} - 2\mathbf{x}_2 \mathbf{x}'_1 \mathbf{M}_{11}^{-1} \mathbf{M}_{12} + \mathbf{x}_2 \mathbf{x}'_2] \boldsymbol{\beta}_2 \, d\mathbf{x}$$

At this point it is convenient to introduce the concept of **region moments**. Letting  $\boldsymbol{\alpha}_2 = \sqrt{N} \boldsymbol{\beta}_2 / \sigma$ , ASB can be written as

$$ASB = \boldsymbol{\alpha}'_2 [\mathbf{M}'_{12} \mathbf{M}_{11}^{-1} \boldsymbol{\mu}_{11} \mathbf{M}_{11}^{-1} \mathbf{M}_{12} - 2\boldsymbol{\mu}'_2 \mathbf{M}_{11}^{-1} \mathbf{M}_{12} + \boldsymbol{\mu}_{22}] \boldsymbol{\alpha}_2 \tag{10.16}$$

where

$$\boldsymbol{\mu}_{11} = K \int_R \mathbf{x}_1 \mathbf{x}'_1 \, d\mathbf{x}, \quad \boldsymbol{\mu}_{22} = K \int_R \mathbf{x}_2 \mathbf{x}'_2 \, d\mathbf{x} \quad \text{and} \quad \boldsymbol{\mu}_{12} = K \int_R \mathbf{x}_1 \mathbf{x}'_2 \, d\mathbf{x}.$$



The quantities  $\mu_{11}$ ,  $\mu_{22}$ , and  $\mu_{12}$  are **region moment matrices**. The matrix  $\mu_{11}$  contains first and second moments. In fact,

$$\mu_{11} = K \int_R \mathbf{x}_1 \mathbf{x}'_1 = K \int_R \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} [1, x_1, x_2, \dots, x_k] d\mathbf{x}$$

Because  $\int_R$  implies  $\int_{-1}^1 \int_{-1}^1 \dots \int_{-1}^1$ , this first- and second-order region moment matrix  $\mu_{11}$  is given by

$$\mu_{11} = \begin{bmatrix} 1 & | & 0 & \dots & 0 \\ \hline 0 & & & & \\ 0 & & & & \\ \vdots & & & & \\ 0 & & & & \frac{1}{3} I_k \end{bmatrix}$$

The matrix  $\mu_{22}$ , in a similar fashion, displays region moments through order four, while  $\mu_{12}$  contains region moments through order three.

An easy prescription for the form of the design that minimizes ASB is obtained by writing the average squared bias in Equation 10.15 as

$$ASB = \alpha'_2 [(\mu_{22} - \mu'_{12} \mu_{11}^{-1} \mu_{12}) + (\mathbf{M}_{11}^{-1} \mathbf{M}_{12} - \mu_{11}^{-1} \mu_{12})' \mu_{11} (\mathbf{M}_{11}^{-1} \mathbf{M}_{12} - \mu_{11}^{-1} \mu_{12})] \alpha_2 \tag{10.17}$$

In the second term in Equation 10.17 the matrix  $\mu_{11}$  is positive definite. [see Box and Draper (1987) and Myers (1976)]. As a result, a sufficient condition for a **minimum bias design** (i.e., a design that minimizes ASB) is to let

$$\mathbf{M}_{11} = \mu_{11} \quad \text{and} \quad \mathbf{M}_{12} = \mu_{12} \tag{10.18}$$

that is, the design moment matrices  $\mathbf{M}_{11}$  and  $\mathbf{M}_{12}$  should be equal to the corresponding region moment matrices  $\mu_{11}$  and  $\mu_{12}$ . As a result, the minimum bias design is one in which  $ASB = \alpha'_2 [\mu_{22} - \mu'_{12} \mu_{11}^{-1} \mu_{12}] \alpha_2$ . Note that selection of an appropriate design reduces to the selection of design moment matrices  $\mathbf{M}_{11}$  and  $\mathbf{M}_{12}$ . In the case where a first-order model is being fitted and one seeks to protect against a second-order model,  $\mathbf{M}_{11}$  contains moments through order two and  $\mathbf{M}_{12}$  contains moments through order three. However, note that Equation 10.17 allows for the minimum ASB in general—that is, if one fits a model with the form  $\hat{\mathbf{y}} = \mathbf{X}_1 \mathbf{b}_1$  and protection is sought against a model  $\mathbf{y} = \mathbf{X}_1 \beta_1 + \mathbf{X}_2 \beta_2$ .

**Single Design Variable—Minimum Bias Design** It is useful to write one of the general expressions in Equation 10.18 for the special case of a single design variable. Here, of course, the design becomes the selection of the spacing of points in the interval  $[-1, +1]$  in the design variable. As a result, we assume a fitted first-order model

$$\hat{y}(x) = b_0 + b_1 x$$

when the true structure is best approximated by

$$E[y(x)] = \beta_0 + \beta_1 x + \beta_{11} x^2$$

We have an experiment that results in

$$\mathbf{X}_1 = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix}$$

with

$$\mathbf{X}_2 = \begin{bmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_N^2 \end{bmatrix}.$$

The relevant moment matrices are

$$\mathbf{M}_{11} = \frac{\mathbf{X}'_1 \mathbf{X}_1}{N} = \begin{bmatrix} 1 & [1] \\ [1] & [11] \end{bmatrix}$$

and

$$\mathbf{M}_{12} = \begin{bmatrix} [11] \\ [111] \end{bmatrix}$$

where  $[11] = (1/N)\sum_u x_u^2$ ,  $[1] = (1/N)\sum_u x_u$ , and  $[111] = (1/N)\sum_u x_u^3$ . Now, the region moment matrices  $\boldsymbol{\mu}_{11}$  and  $\boldsymbol{\mu}_{12}$  are of the same dimension and structure as their design moment counterparts  $\mathbf{M}_{11}$  and  $\mathbf{M}_{12}$ . In fact,

$$\boldsymbol{\mu}_{11} = \frac{1}{2} \int_{-1}^1 \begin{bmatrix} 1 \\ x \end{bmatrix} [1 \quad x] dx = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$$

and

$$\boldsymbol{\mu}_{12} = \frac{1}{2} \int_{-1}^1 \begin{bmatrix} 1 \\ x \end{bmatrix} [x^2] dx = \begin{bmatrix} \frac{1}{3} \\ 0 \end{bmatrix}$$

For the interval  $[-1, +1]$ , which represents the region of interest in  $x$ , the odd region moments are all zero. The quantity  $\frac{1}{3}$  in  $\mu_{11}$  and  $\mu_{12}$  is the value of the second pure region moment  $\frac{1}{2} \int_{-1}^1 x^2 dx$ . As a result, a very simple application of the results in Equation 10.18 suggests:

A *minimum bias* (minimum ASB) design for a fitted first-order model with protection against a model of second order can be achieved by having  $[1]=0$ ,  $[11]=\frac{1}{3}$  and  $[111]=0$ .

The above result is not at all counterintuitive and, in fact, is certainly the kind of result one expects from the discussion in Section 10.1. Clearly the **minimum variance design** (i.e., the design that minimizes APV for this case) will have  $[1]=0$  and  $[11]=1.0$  (all points pushed to the  $\pm 1$  levels). The minimum bias design requires levels to be placed inward from the  $\pm 1$  extremes.

For the case of the single design variable, there are many designs that meet the moment conditions that result in minimum bias—that is, minimum average squared bias. The requirement of zero first and third moments requires only an equal sample size design with levels spaced symmetrically around the center of the design interval. It is interesting, then, that the design that best protects against the quadratic model can take many forms. For example, a *two-level design* with, say,  $N$  observations ( $N$  even) requires

$$N/2 \text{ observations at } x = \sqrt{\frac{1}{3}} = 0.58$$

$$N/2 \text{ observations at } x = -\sqrt{\frac{1}{3}} = -0.58$$

Of course there are obvious advantages in using more than two levels. The extra degrees of freedom for lack of fit are certainly useful. Thus, for three levels and  $N/3$  runs at each level, the design is given by

$$\frac{N}{3} \text{ at } x = a, \quad \frac{N}{3} \text{ at } x = 0, \quad \frac{N}{3} \text{ at } x = -a$$

where, to allow  $[11]=\frac{1}{3}$ , we have

$$\frac{2}{3}a^2 = \frac{1}{3}$$

and thus

$$a = \sqrt{\frac{1}{2}} \cong 0.707$$

This demonstrates the considerable flexibility in constructing the minimum bias design. One can have various numbers of center runs and divide the remaining runs between two levels symmetric around  $x=0$ . For example, suppose 15 runs are allowed. Suppose, also,

that three runs are to be used in the center. The remaining 12 runs are then evenly divided between  $-a$  and  $a$ , where

$$\left(\frac{4}{5}\right) a^2 = \frac{1}{3}$$

and thus

$$a = \sqrt{\frac{5}{12}} \cong 0.644$$

Thus, it is important to be sure that all three moment conditions hold. The minimum bias design need not have equally spaced levels. For example, consider the design with levels  $-a_2, -a_1, a_1,$  and  $a_2$  (no center runs) with  $N/4$  runs at each level. The values  $a_1$  and  $a_2$  are such that

$$\frac{a_1^2 + a_2^2}{2} = \frac{1}{3}$$

There are many choices for  $a_1$  and  $a_2$  and thus many four-level designs. One example is  $a_1 = \sqrt{\frac{1}{6}}$  and  $a_2 = \sqrt{\frac{1}{2}}$  and another is  $a_1 = \frac{1}{2}$  and  $a_2 = \sqrt{\frac{5}{12}}$ .

**Multiple Design Variables—Minimum Bias Design** The use of Equation 10.18 is general and applies in the case of the fitted model

$$\hat{y}(\mathbf{x}) = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k$$

Suppose we wish to protect against a full second-order model characterized by  $\mathbf{X}_2$  in Equation 10.15. Suppose the region of interest is a cuboidal region with  $-1 < x_i \leq 1$  for  $i = 1, 2, \dots, k$ . The pertinent region moments are

$$\mu_{11} = \left[ \begin{array}{c|cc} 1 & 0 & \dots & 0 \\ \hline 0 & & \frac{1}{3}I_k & \\ \vdots & & & \\ 0 & & & \end{array} \right]$$

$$\mu_{12} = \left[ \begin{array}{c|cccc} \frac{1}{3} & \frac{1}{3} & \dots & \frac{1}{3} & 0 & 0 & \dots & 0 \\ \hline & & & & \mathbf{0} & & & \end{array} \right]$$

As a result, if we again accomplish minimum ASB by equating design moments to region moments through order three, we require  $\mathbf{M}_{11} = \mu_{11}$  and  $\mathbf{M}_{12} = \mu_{12}$ . This results in setting

all odd moments through order three equal to zero. It should be noted that

$$\boldsymbol{\mu}_{12} = \frac{1}{K} \int_R \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} [x_1^2, x_2^2, \dots, x_k^2, x_1x_2, \dots, x_kx_{k-1}] d\mathbf{x}$$

and thus  $\boldsymbol{\mu}_{12}$  is a  $k + 1$  by  $k + k(k - 1)/2$  matrix containing second and third moments. The moment matrix  $\mathbf{M}_{12}$  has the same dimension and structure. As a result, the minimum ASB can be obtained by setting

$$\begin{aligned} [i] &= 0 & i = 1, 2, \dots, k \\ [ii] &= \frac{1}{3}, & i = 1, 2, \dots, k \\ [iii] &= 0, & i = 1, 2, \dots, k \\ [ij] &= 0, & i \neq j \\ [ij] &= 0, & i \neq j \\ [ijk] &= 0, & i \neq j \neq k \end{aligned} \tag{10.19}$$

The two-level full factorial design in  $k$  variables contains the type of symmetries that result in the above odd moments being zero. In order to achieve a design with each second pure moment equal to  $\frac{1}{3}$ , the level of the factorial points must be  $\pm g$ , where  $g = \sqrt{\frac{1}{3}} = 0.58$ .

**What About Fractional Factorials and Center Runs?** It is clear that complete  $2^k$  factorials with levels brought in considerably from the  $\pm 1$  extremes in the case of a cuboidal region satisfy the moment conditions given in Equation 10.19 for a minimum bias design. But will fractional factorials suffice? For the case of odd design moments, the regular  $2^{k-p}$  fractional factorial of resolution  $\geq III$ , or even the Plackett–Burman designs, produce moments  $[i]$ ,  $[iii]$ ,  $[ij]$ , and  $[iij]$  that are zero. However, the resolution III design *does not result in*  $[ijk] = 0$ . This should be obvious, because all  $x_i$  will not be orthogonal to all  $x_jx_k$  ( $i \neq j \neq k$ ) for the case of resolution III. However, the following statement can be made:

A minimum bias design in a cuboidal region for a first-order model with protection against a model of second order is achieved with a  $2^k$  factorial design or a  $2^{k-p}$  fraction with resolution  $\geq IV$ . If the extremes of the cuboidal region are scaled to  $\pm 1$ , the design levels will be at  $\pm 0.58$ .

Center runs are very useful when one desires to afford protection against curvature. The conditions in Equation 10.19 on the odd moments are not altered by the use of center runs, and the condition on the second pure moment is achieved with the intuitively appealing result that the design levels can be moved closer to the  $\pm 1$  extremes. This will likely be an attractive alternative. For example, suppose one uses a  $\frac{1}{2}$  fraction of a  $2^4$  factorial with

$n_0 = 4$  center runs. The extremes for the region of interest are scaled to  $\pm 1$ , and a minimum bias design is given by

$$\mathbf{D} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ -g & -g & -g & -g \\ g & g & -g & -g \\ g & -g & g & -g \\ g & -g & -g & g \\ -g & g & g & -g \\ -g & g & -g & g \\ -g & -g & g & g \\ g & g & g & g \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

with  $g$  determined so that  $[ii] = \frac{1}{3}$ . In fact,

$$g = \sqrt{\left(\frac{12}{8}\right)\left(\frac{1}{3}\right)} = 0.707$$

As a second example consider a  $\frac{1}{2}$  fraction of a  $2^5$  factorial with  $n_c = 8$  center runs. Using  $\pm g$  notation to indicate design levels, a minimum bias design is produced for a cuboidal region of interest scaled to  $\pm 1$  by using a resolution V fraction with

$$g = \sqrt{\left(\frac{24}{16}\right)\left(\frac{1}{3}\right)} = 0.707$$

It should be clear that the selection of the region of interest is important. Minimizing the bias due to an underspecification of the model requires that points be placed inward from the extremes of that region, but the criterion (i.e., squared bias) is integrated, or **averaged**, over the entire region. Anderson-Cook et al. (2009) consider several examples of comparing designs on cuboidal regions for protecting against higher order models. While the cuboidal region of interest is certainly quite common in practice, there are still instances in which a spherical region is the operative region of interest. In the next section we discuss the protection against bias due to the presence of a second-order model when the region of interest is spherical.

### 10.2.2 Minimum Bias Designs for a Spherical Region of Interest

Consider again a fitted first-order model with protection against a second-order model, but the squared bias in prediction is average over a **spherical region**. That is, there are ranges in the variables that aid in structuring the region of interest, but locations that represent simultaneous extremes in the variables are not included. As we discussed at length in

Chapter 8, this condition suggests a region not unlike a sphere. For simplicity in what follows, we will assume that the region of interest is a unit sphere—that is, the region

$$\sum_{i=1}^k x_i^2 \leq 1 \tag{10.20}$$

Thus the variables are centered so that the center of the region is  $(0, 0, \dots, 0)$  and the radius of the sphere is 1.0.

As in the case of the cuboidal region, the minimum bias design for the spherical region can be obtained by achieving moment conditions that are governed by Equation 10.18. Again this involves equating design moments to region moments through order three. For the region given by Equation 10.20, all odd region moments are zero and the second pure region moment is given by  $K \int_R x_i^2 d\mathbf{x} = 1/(k + 2)$ . Notice that for the special case in which  $k = 1$ , the value is  $\frac{1}{3}$ . The minimum bias design for the spherical region of Equation 10.20 is achieved by a design with moment conditions:

$$\begin{aligned} [i] &= [ij] = [iii] = [ijj] = [ijk] = 0 && (i \neq j \neq k) \\ [ii] &= \frac{1}{k + 2} && (i = 1, 2, \dots, k) \end{aligned} \tag{10.21}$$

Again it becomes clear that the points that produce the minimum bias design are set in from the perimeter of the sphere. In addition, the symmetry of the region, like that of the cube, requires a design symmetry that forces odd design moments through order three to be zero. Once again the two-level factorial meets all design requirements, with the  $[ii] = 1/(k + 2)$  produced by the choice of the specific factorial levels  $\pm g$ .

**Example 10.1** It is important to understand the importance of the scaling specified by the region  $\sum_{i=1}^k x_i^2 = 1.0$ . This implies that the factorial points that are placed at the perimeter of the region reside at  $\pm 1/\sqrt{k}$ . The points  $\pm 1$  are outside the region of interest in the metric invoked by the region. Suppose then that  $k = 4$  and a  $2^4$  factorial with  $n_c = 5$  center runs is employed. The zero odd moments satisfy the requirements of a minimum bias design according to Equation 10.21. The second pure moment  $[ii]$  should be

$$[ii] = \frac{1}{k + 2} = \frac{1}{6}$$

As a result, with  $n_c = 5$  center runs, the  $2^4$  factorial contains the design levels  $\pm g$  at

$$\begin{aligned} g &= \sqrt{\left(\frac{1}{6}\right) \frac{21}{16}} \\ &= \sqrt{0.21} \\ &\cong 0.46 \end{aligned}$$

Therefore, the minimum bias design requires factorial points that are placed at 0.46, whereas the factorial points at the perimeter of the design region are scaled to  $1/\sqrt{4} = 0.5$ .

**10.2.3 Simultaneous Consideration of Bias and Variance**

Early in Section 10.2 we introduced the notion of average mean squared error (AMSE) and the partition of Equation 10.11 into APV and ASB. From that point on we have focused strictly on bias and thus the minimization of ASB. At this point let us consider the use of the entire AMSE. From Equations 10.11 and 10.16 we have

$$\begin{aligned}
 \text{AMSE} &= \alpha_2' [(\mu_{22} - \mu_{12}' \mu_{11}^{-1} \mu_{12}) \\
 &\quad + (\mathbf{M}_{11}^{-1} \mathbf{M}_{12} - \mu_{11}^{-1} \mu_{12})' \mu_{11} (\mathbf{M}_{11}^{-1} \mathbf{M}_{12} - \mu_{11}^{-1} \mu_{12})] \alpha_2 \\
 &\quad + \frac{NK}{\sigma^2} \int_R \text{Var}[\hat{y}(\mathbf{x})] d\mathbf{x}
 \end{aligned}
 \tag{10.22}$$

As before, we are assuming, in this general formulation, that we are fitting a model  $\hat{y} = \mathbf{x}'_1 \mathbf{b}_1$  and the true model is given by  $E(y) = \mathbf{x}'_1 \beta_1 + \mathbf{x}'_2 \beta_2$ . Of course our emphasis focuses on  $\mathbf{x}'_1 \beta_1$  representing a first-order model, while  $\mathbf{x}'_2 \beta_2$  contains second-order terms. The average prediction variance can be written

$$\frac{NK}{\sigma^2} \int_R \text{Var}[\hat{y}(\mathbf{x})] d\mathbf{x} = K \int_R \mathbf{x}'_1 \mathbf{M}_{11}^{-1} \mathbf{x}_1 d\mathbf{x} = \text{tr}[\mathbf{M}_{11}^{-1} \mu_{11}]
 \tag{10.23}$$

Clearly a desirable design criterion is to select the design that minimizes AMSE. AMSE is a function of design moments, region moments, and  $\alpha_2$ , a standardized vector of parameters. It was clear in the previous section that the bias component can be minimized in spite of  $\alpha_2$ . However, *AMSE cannot be minimized with respect to design moments without knowledge of  $\alpha_2$ .*

As a result, the practical use of AMSE is difficult at best. However, it is of interest to determine what role bias plays in the AMSE and, in particular, how effective the minimum bias design is in controlling AMSE. The following subsection deals with these matters and discusses the use of compromise designs—that is, designs that are not minimum AMSE but are approximately minimum AMSE.

**10.2.4 How Important Is Bias?**

Let us consider the relative performance of the minimum variance and minimum bias designs for  $k = 1$  for various magnitudes of bias. In other words, let us ask which design is able to more closely emulate the performance of the optimal design; that is, the design that minimizes AMSE. In Table 10.1, various degrees of bias are listed in the form of  $\sqrt{N} \beta_{11} / \sigma$ , the standardized regression coefficient of the quadratic term. Here we are assuming a fitted model

$$\hat{y} = b_0 + b_1 x$$

and a true model

$$E(y) = \beta_0 + \beta_1 x + \beta_{11} x^2$$



**TABLE 10.1** Values of Optimal [11] and AMSE for Optimal Design, Minimum Variance Design, and Minimum Bias Design

| $\sqrt{N}\beta_{11}/\sigma$ | APV/ASB | [11]<br>(Optimum) | AMSE<br>(Optimum) | AMSE<br>([11] = 1.0) | AMSE<br>([11] = $\frac{1}{3}$ ) |
|-----------------------------|---------|-------------------|-------------------|----------------------|---------------------------------|
| 0                           |         | 1.0               | 1.333             | 1.333                | 2.0                             |
| 0.50                        | 10      | 0.999             | 1.466             | 1.466                | 2.022                           |
| 1.0                         | 7       | 0.687             | 1.699             | 1.867                | 2.088                           |
| 1.5                         | 5       | 0.565             | 1.910             | 2.533                | 2.205                           |
| 2.0                         | 3.7     | 0.500             | 2.133             | 3.466                | 2.352                           |
| 2.5                         | 3       | 0.462             | 2.382             | 4.672                | 2.551                           |
| 3.0                         | 2       | 0.432             | 2.666             | 6.135                | 2.805                           |
| 4.0                         | 1.5     | 0.399             | 3.333             | 9.872                | 3.422                           |
| 4.5                         | 1.0     | 0.388             | 3.718             | 11.521               | 3.798                           |
| 5.0                         | 0.9     | 0.379             | 4.153             | 14.677               | 4.222                           |
| 7.0                         | 0.4     | 0.359             | 6.316             | 27.465               | 6.355                           |

with  $R$  scaled to  $[-1, +1]$ . For each value of  $\sqrt{N}\beta_{11}/\sigma$ , the optimal design (minimizing AMSE) was computed and thus the minimum value of AMSE was also computed. In addition, the AMSE for the minimum bias ( $[11] = \frac{1}{3}$ ) design and the AMSE for the minimum variance design ( $[11] = 1.0$ ) were computed. These quantities are listed as well as  $\sqrt{N}\beta_{11}/\sigma$ . The degree of bias is quantified by the ratio APV/ASB—that is, the ratio of the variance portion of the AMSE to the bias contribution in the case of the optimal design.

Study of Table 10.1 makes it clear that bias is an important consideration. Even when the bias makes a relatively small contribution—for example, when APV/ASB = 5 or APV/ASB = 3.7—the minimum bias design results in a smaller AMSE than that which results from the minimum variance design. In addition, it is clear that the minimum bias approach is considerably more robust than the minimum variance design. The minimum bias design is never far from optimal (in terms of mean squared error), whereas the minimum variance design performs quite poorly when the bias present is such that APV/ASB < 5.

Table 10.1 suggests that perhaps a compromise design for which [11] is slightly larger than the minimum bias value of  $\frac{1}{3}$  might be appropriate, because it should be extremely robust—that is, very close in AMSE to that of the optimal design. A design with a second moment [11] of 0.40–0.45 is in fact extremely robust. An impression can be obtained from Table 10.2, which shows the performance of the design with  $[11] = 0.40$ . Obviously, the use of  $[11] = 0.45$  or  $0.5$  would produce even better performance for very small values of the bias parameter  $\sqrt{N}\beta_{11}/\sigma$ .

Tables similar to Tables 10.1 and 10.2 for  $k > 1$  design variables can be constructed. The results teach essentially the same lesson: if there is any suspicion at all that there is curvature in the system, the minimum variance design is very likely to be counterproductive. In fact, a reasonable rule of thumb to allow for possible model bias is to use a design with zero moments through order three and second pure moments  $[ii]$  that are 20–25% higher than that suggested by the minimum bias design. This design prescription serves as a reasonable compromise and will provide a design that will not be far away in performance from the optimal (minimum mean squared error) design over a wide range of bias in the system.

**TABLE 10.2** AMSE Values for the Optimal Design and Compromise Design of [11] = 0.4

| $\sqrt{N}\beta_{11}/\sigma$ | AMSE (Optimal) | AMSE ([11] = 0.40) |
|-----------------------------|----------------|--------------------|
| 0                           | 1.33           | 1.833              |
| 0.5                         | 1.467          | 1.86               |
| 1.0                         | 1.699          | 1.927              |
| 1.5                         | 1.910          | 2.043              |
| 2.0                         | 2.133          | 2.206              |
| 2.5                         | 2.382          | 2.416              |
| 3.0                         | 2.666          | 2.67               |
| 4.0                         | 3.333          | 3.327              |
| 5.0                         | 4.153          | 4.167              |
| 7.0                         | 6.316          | 6.406              |

The above design moment recommendations can sometimes be a bit difficult to implement in practice. The problem often stems from uncertainty regarding what is truly the region of interest. Even in the case of what might be perceived to be a cuboidal region, the exact locations, in the design variables, of the corners of the region are not always known. Certainly in preliminary studies where variable screening is involved, choices of ranges of variables often require much thought. As a result, the recommendations made here should be confined to situations in which the researcher is not engaged in variable screening but is reasonably certain regarding the relative importance of design variables and the region in which prediction or predicted values are sought. Thus, of course, protection is achieved by designing an experiment that has resolution at least IV, and the design levels are brought in from the perimeter of the region in which one is required to predict response.

### 10.3 ERRORS IN CONTROL OF DESIGN LEVELS

In practical situations one plans to design experiments according to some master plan that entails the choice of a combination of design levels and a random assignment of these design levels to experimental units. However, it is often difficult to *control* design levels. This is true even in laboratory situations. This suggests a need to determine how one can assess the effects of **errors in control** of design levels. For an excellent account see Box (1963). Recall that this is one of the 11 desirable properties of response surface designs described in Section 8.1.

The approach to assessment of errors in design levels in an RSM setting depends on the structure of the error. In some (perhaps most) instances there is a chosen level according to the design plan. This **aimed at** level is reported and used in the analysis even though it is known that it is subject to error. The actual level that is experienced by the process is not known. We call this the **fixed design level** case.

**Fixed Design Level Case** One must keep in mind that the actual levels are not observed in this case. The deviation between the actual and the planned level is a random variable. That is, for the  $u$ th run of the  $j$ th design variable we have

$$w_{ju} = x_{ju} + \theta_{ju} \quad (u = 1, 2, \dots, N, j = 1, 2, \dots, k) \quad (10.24)$$

where  $w_{ju}$  is the actual level,  $x_{ju}$  is the fixed and planned level, and  $\theta_{ju}$  is a random error in control. It may be assumed that

$$E(\theta_{ju}) = 0$$

$$\text{Var}(\theta_{ju}) = \sigma_j^2 \quad (u = 1, 2, \dots, N, \quad j = 1, 2, \dots, k)$$

For the case of multiple design levels, the variances of,  $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$  are the variances of the errors in control.

**First-Order Model** Consider the fixed level case when one fits a first-order model. The model, in terms of the true (yet unknown) levels, is given by

$$y = \beta_0 + \beta_1 w_1 + \beta_2 w_2 + \dots + \beta_k w_k + \varepsilon \tag{10.25}$$

Because the analysis will be done on the planned levels, it is instructive to write the model in Equation 10.25 as

$$y = \beta_0 + \beta_1(x_1 + \theta_1) + \beta_2(x_2 + \theta_2) + \dots + \beta_k(x_k + \theta_k) + \varepsilon$$

$$= \beta_0 + \sum_{j=1}^k \beta_j x_j + \left\{ \varepsilon + \sum_{j=1}^k \beta_j \theta_j \right\}$$

$$= \beta_0 + \sum_{j=1}^k \beta_j x_j + \varepsilon^* \tag{10.26}$$

The  $\theta_j$  are random variables as described by Equation 10.24, but the  $x_j$  are fixed (pre-determined by the design choice). As a result, Equation 10.26 represents an almost standard RSM model with model error

$$\varepsilon^* = \varepsilon + \sum_{j=1}^k \beta_j \theta_j \tag{10.27}$$

It becomes obvious from Equation 10.27 that the errors associated with controlling design variables are transmitted to the model error and the result is an inflated error variance. As a result, if the model in Equation 10.26 is used and least squares is used to estimate  $\beta_0, \beta_1, \beta_2, \dots, \beta_k$  with ordinary least squares, we have the following:

1. The regression coefficients are unbiased.
2. The variance-covariance matrix of  $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$  is given by  $(\mathbf{X}'\mathbf{X})^{-1}\sigma^{*2}$ , where  $\sigma^{*2} = \sigma^2 + \sum_{j=1}^k \beta_j^2 \sigma_j^2$ , assuming that the  $\theta_j$  are independent. This results in inflated error variances. Depending on the magnitude of the  $\sigma_j^2$ , the inflation may be severe.
3. Standard RSM (e.g., variable screening) and steepest ascent are valid, although they may be rendered considerably less effective if the  $\sigma_j^2$  are large. Of course, if the  $\sigma_j^2$  are extremely small compared to the model variance  $\sigma^2$ , these errors will not result in major difficulties.

We emphasized that the conclusions drawn in 1 and 2 above do depend on the assumptions made regarding the  $\theta_j$ . Of particular interest is the assumption that the  $\theta_j$  are unbiased—that is,  $E(\theta_j) = 0$ . This assumption is certainly a reasonable one in most cases. However, if it does not hold, the least squares estimator of  $\beta_0$  becomes biased.

**Second-Order Model** In the case of the second-order model, errors in control result in more difficulties than one experiences in the first-order case if the error variances are not negligible. Consider again Equation 10.24 and the second-order model given by

$$\begin{aligned}
 y &= \beta_0 + \sum_{j=1}^k \beta_j w_j + \sum_{j=1}^k \beta_{jj} w_j^2 + \sum_{i < j=2}^k \sum_{i < j=2}^k \beta_{ij} w_i w_j + \epsilon \\
 &= \beta_0 + \sum_{j=1}^k \beta_j (x_j + \theta_j) + \sum_{j=1}^k \beta_{jj} (x_j + \theta_j)^2 + \sum_{i < j=2}^k \sum_{i < j=2}^k \beta_{ij} (x_i + \theta_i)(x_j + \theta_j) + \epsilon \\
 &= \beta_0 + \sum_{j=1}^k \beta_j x_j + \sum_{j=1}^k \beta_{jj} x_j^2 + \sum_{i < j=2}^k \sum_{i < j=2}^k \beta_{ij} x_i x_j \\
 &\quad + \left\{ \epsilon + \sum_{j=1}^k \beta_j \theta_j + \sum_{j=1}^k \beta_{jj} \theta_j (2x_j + \theta_j) + \sum_{i < j=2}^k \sum_{i < j=2}^k \beta_{ij} [\theta_i x_j + x_i \theta_j + \theta_i \theta_j] \right\} \\
 &= \beta_0 + \sum_{j=1}^k \beta_j x_j + \sum_{j=1}^k \beta_{jj} x_j^2 + \sum_{i < j=2}^k \sum_{i < j=2}^k \beta_{ij} x_i x_j + \epsilon^* \tag{10.28}
 \end{aligned}$$

The model in Equation 10.28 is best analyzed by considering the expectation and variance of  $\epsilon^*$ . Once again we have a standard RSM model with error  $\epsilon^*$ . However, in this case the properties of  $\epsilon^*$  are more complicated. The expected value of  $\epsilon^*$  is given by

$$E(\epsilon^*) = \sum_{j=1}^k \beta_{jj} \sigma_j^2$$

Because  $E(\epsilon^*) \neq 0$  and the **error bias**  $\sum_{j=1}^k \beta_{jj} \sigma_j^2$  is independent of the  $x_j$ , the least squares estimator of  $\beta_0$  is biased by  $\sum_{j=1}^k \beta_{jj} \sigma_j^2$ .

In order to better illustrate the result in the case of the second-order model, we rewrite  $\epsilon^*$  in Equation 10.28 as

$$\begin{aligned}
 \epsilon^* &= \epsilon + \sum_{i < j=2}^k \sum_{i < j=2}^k \theta_i (\beta_{ij} x_j) + \sum_{i < j=2}^k \sum_{i < j=2}^k \theta_j \beta_{ij} x_i + \sum_{i < j=2}^k \sum_{i < j=2}^k \beta_{ij} \theta_i \theta_j \\
 &\quad + \sum_{j=1}^k \theta_j \beta_j + 2 \sum_{j=1}^k \beta_{jj} \theta_j x_j + \sum_{j=1}^k \beta_{jj} \theta_j^2 \tag{10.29}
 \end{aligned}$$

If one considers the variance of  $\varepsilon^*$  in Equation 10.29, it becomes apparent that:

1. The error variance is inflated by errors in control transmitted to  $\varepsilon^*$  just as in the first-order case.
2. The error variance will no longer be homogeneous. Rather,  $\text{Var}(\varepsilon^*)$  depends on design levels.

While we will not display  $\sigma^{*2} = \text{Var} \varepsilon^*$ , it should be apparent that terms such as  $\sum_{j=1}^k \beta_j^2 \sigma_j^2$ ,  $\sum_{j=1}^k \beta_{ij}^2 [E(\theta_j^4) - \sigma_j^4]$ , and  $\sum \sum_{i<1} \beta_{ij}^2 \sigma_i^2 \sigma_j^2$  inflate the error variance, whereas terms such as  $4 \sum_{j=1}^k \beta_{ij}^2 x_j^2 \sigma_j^2$ ,  $\sum \sum_{i<j} \beta_{ij}^2 x_j^2 \sigma_i^2$ , and  $\sum \sum_{i<j} \beta_{ij}^2 x_i^2 \sigma_j^2$  render the error variance nonhomogeneous.

**Further Comments** The degree of difficulty caused by errors in control obviously depends on the model complexity and on the size of the error variances  $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$ . If the latter are dominated by the model error variance  $\sigma^2$ , then the effect of errors in control is negligible.

## 10.4 EXPERIMENTS WITH COMPUTER MODELS

We customarily think of applying RSM to a physical process, such as chemical vapor deposition in semiconductor manufacturing, wave soldering, or machining. However, RSM can also be successfully applied to computer simulation models of physical systems. Computer models are becoming increasingly common, and they can be used as proxies for many complex processes that are difficult or expensive to manipulate. In some applications, there are restrictions on what conditions can be explored with a physical experiment because of cost, safety, or regulations. In other cases, having a computer model (or code) allows for much more rapid exploration of alternatives, development of prototypes and new products—increasing competitive advantage and speed to market. Computer models can often have a large number of inputs and result in multiple responses, which can be either scalar (a single value) or functional (a collection of values connected over time or space). Hence design of experiments has an important role to play in the selection of a preferred set of input combinations to explore.

In such computer modeling applications, the role of RSM is different, as the data obtained from runs of the computer model can be used to build a model of the system being modeled by the computer simulation, called a **metamodel or emulator**. Since for some computer models, obtaining one observation from the code may take considerably computer runtime, understanding about the characteristics of the system can be gleaned from exploring the estimated model and optimization is carried out on the metamodel. The assumption is that if the computer simulation model is a faithful representation of the real system, then the RSM optimization will result in adequate determination of the optimum conditions for the real system.

Generally, there are two types of simulation models, **stochastic** and **deterministic**. In a stochastic simulation model, the output responses are random variables. Examples include (a) systems simulations such as the factory planning and scheduling models used in the semiconductor industry and traffic flow simulators employed by civil engineers and (b) Monte Carlo simulations that sample from probability distributions to study complex

mathematical phenomena that do not have direct analytical solutions. See Fang and Sudjianto (2006) for more discussion.

In deterministic simulation models, the output responses are not random variables; they are entirely deterministic quantities whose values are determined by the (often highly complex) mathematical models upon which the computer model is based. Deterministic simulation models are often used by engineers and scientists as computer-based design tools. Typical examples are circuit simulators used for designing electronic circuits and semiconductor devices, finite element analysis models for mechanical and structural design, and computational models for physical phenomena such as fluid dynamics. These are often very complex models, requiring considerable computer resources and time to run.

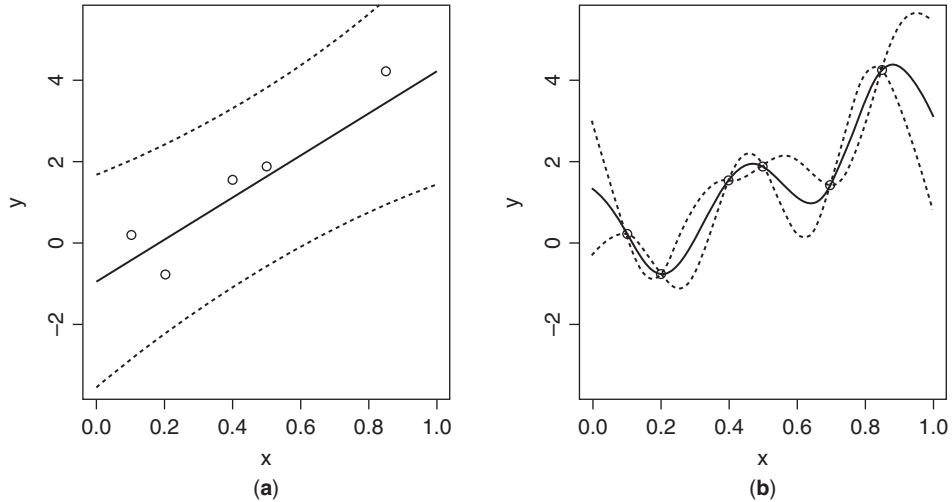
Before considering what designs and analyses to use when studying computer models, it is helpful to compare data obtained from physical experiments versus from computer models. Data from both types of experiments can be expensive: for physical experiments, the cost comes from set-up and running the experiment as well as measuring the response values. For computer experiments, the development of the codes is labor- and time-intensive and often requires subject matter expertise. But here we assume that the computer model is already available and we wish to use it to explore the underlying relationship of interest. Obtaining the data itself is also often expensive. Because of the complexity of the code, it may require substantial amounts of computer power and runtime to obtain results even for a moderate number of input combinations. Hence, for both types of experiments, the ability to obtain a good estimate of the relationship efficiently from small to moderate amounts of data is important.

A key difference between physical and computer experiments is the range over which experimentation may take place. For physical experiments, the goal is often to restrict the region of interest to a local region where the relationship between inputs and response are well characterized by a low-order polynomial. However, for computer experiments, an experiment might be sought to characterize the relationship over much larger regions of the input space, since there are not the physical limitations often imposed by where admissible values may be obtained in the region of operability. The boundaries for where observations can be obtained from a computer code are often dictated by changing science or engineering mechanisms that have been built into the code.

Data from physical experiments represents observations from the actual process and typically are thought to be unbiased, conditional on having a good measurement device to measure the response. However, because of imperfect measurement devices, small differences in how the input settings might be set, and the natural variability of the process, we typically do not expect to observe identical values for replicate runs of the experiment. In this way, our statistical models for physical processes are defined with an error term to capture and estimate these differences.

On the other hand, data from computer models is the result of a man-made characterization of the relationship, based on the best available science and engineering of the mechanisms driving the process. Depending on the maturity of the code and the depth of underlying knowledge, these codes can range from very accurate to just coarse representations of the main mechanisms. As a result, the user should remain aware that the results from any experiment to explore a computer model are only as good as the underlying knowledge that was available to build the code. Hence, bias or systematic differences between the code and the true process that it is intended to represent are possible.

If the computer model is deterministic, then replicated runs with the same inputs will result in identical values, making this feature undesirable for this type of computer



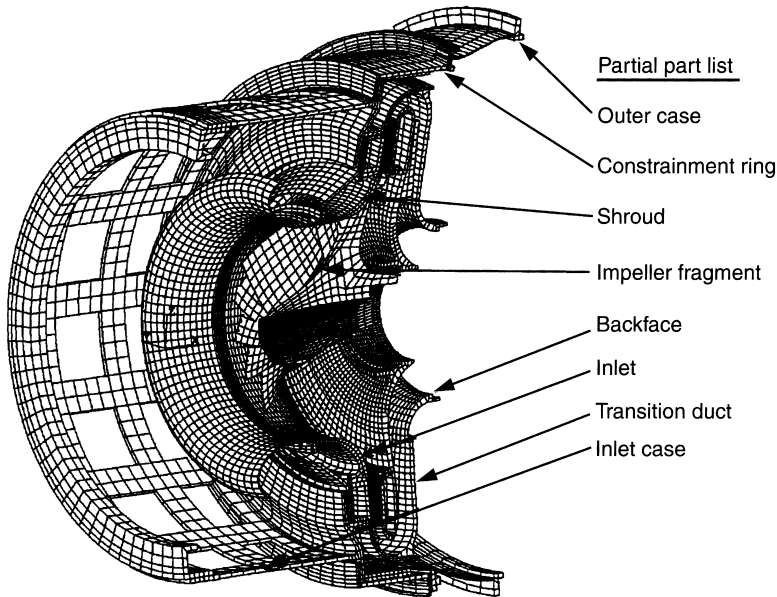
**Figure 10.3** Estimated models with associated uncertainty for data from (a) a physical experiment and (b) from a deterministic computer experiment.

experiments. In these cases, it is helpful to have any estimated metamodel or emulator for the code interpolate between observed points. This makes sense since we believe that the results from the code represent the best available knowledge of the underlying process, and we want to use this information directly. The goal of the emulator is to allow for estimation of other locations in the input space that have not been directly evaluated. If the computer model is stochastic, then there may be some value in obtaining replicates to understand the natural variability.

Another key aspect to consider when comparing physical experiments and computer models is the nature of the underlying relationship being characterized. As we have said throughout previous chapters, RSM is predicated on the assumption that many physical systems are smooth and continuous and can be well approximated, at least in the region of interest, by low-order polynomials. This assumption might not be appropriate for many complex computer codes. For example, sometimes polynomials of higher-order than the usual quadratic response surface models are used. Johnson et al. (2010) compare the performance of different designs when analysis uses higher-order polynomials. Because different options are sometimes needed to describe the underlying relationship, strategies for designed experiments based on different models may require some specialized techniques.

Figure 10.3 illustrates some of the differences between the data and models from physical and computer experiments. Both Figs. 10.3a and 10.3b show the same data for a very simple situation where we are estimating the response from a single input factor. Because the underlying assumptions about the nature of the data differ, this leads to different choices for models and the estimation with associated uncertainty. For data from a physical experiment in Fig. 10.3a, the assumed model is a linear regression line with estimated intercept and slope. Since the observations do not lie directly on the best-fitting line, the magnitude of each miss is used to estimate the natural variation in the data.

If, on the other hand, the data come from a deterministic computer experiment, as shown in Fig. 10.3b, then the selected model is an interpolator, with the best estimated value for an input value with an observed data point matches what was obtained from the computer code.



**Figure 10.4** Finite element model for compressor containment analysis of a turbine engine and partial parts list.

Since we have an exact value for the computer code results at that location, the uncertainty for the emulator is zero at that point, and the uncertainty bands around the function shrink to zero. Note that the uncertainty for the estimated curve increases as we move away from any observation. As a result, a strategy to minimize the worst-case prediction variance of a deterministic computer code anywhere in the design space is to try to have points as spread out as possible throughout the design space. This will minimize the distance between any new location where we wish to predict and an observed observation.

As an example of a situation where a finite element analysis model may be employed, consider the problem of designing a turbine engine to contain a compressor rotor if it fails. Many factors may influence the design, such as engine operating conditions as well as the location, size, and material properties of surrounding parts. Figure 10.4 shows a cutaway view of a typical compressor containment model. Many parameters for each component are potentially important. The thickness, material type, and geometric feature (bend radius, bolt hole size and location, stiffening ribs or gussets, etc.) are engineering design parameters and, potentially, experimental factors that could be included in a response surface model.

One can see that large numbers of factors are potentially important in the design of such a product. Furthermore, the sign or direction of the effect of many of these factors is unknown. For instance, setting factors that increase the axial stiffness of a backface (such as increasing the thickness of the transition duct) may help align a rotor fragment, centering the impact on the containment structure. On the other hand, the increased stiffness may nudge the fragment too much, causing it to miss the primary containment structure. From experience, the design engineers may confidently assume that only a small number of these potentially important factors have a significant effect on the performance of the design in containing a failed part. With computer experiments, this assumption of sparsity may be



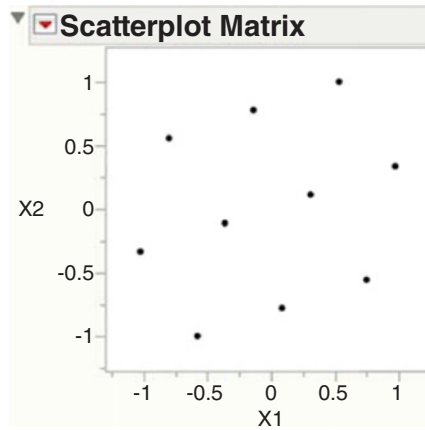
more local: Namely, in some region of the design space, only a small number of factors are actively influencing the response. However, in a computer experiment, it may be that as we move throughout the design space, different subsets of the input factors play a role in influencing the response. This is different from a physical experiment, where in the region of experimentation it is often reasonable to assume that the same subset of factors are active. Detailed analysis or testing of the turbine engine is needed to understand which factors are important and to quantify their effect on the design.

The cost of building a prototype turbine engine is frequently in excess of a million dollars, so studying the effects of these factors using a computer model is very attractive. The type of model used is called a finite element analysis model. Simulating a containment event with a finite element analysis model is very computationally intensive. The model shown in Fig. 10.4 has over 100,000 elements and takes about 90 hr of computer time to model 2 ms of event time. Frequently as much as 10 ms of event time must be modeled. Clearly the need to limit experimentation or simulation is great. Therefore the typical RSM approach of factor screening followed by optimization might well be applied to this scenario to allow some streamlining of the exploration.

#### 10.4.1 Design for Computer Experiments

The choice of design for a computer simulation experiment presents some interesting alternatives. If the experimenter is considering a polynomial model to describe the underlying relationship, then an optimal design such as a  $D$ -optimal or  $I$ -optimal design for the specified model is a possible choice. In recent years, various types of space-filling designs have been suggested for computer experiments. There are several reasons why space-filling designs are thought to be particularly appropriate for deterministic computer models:

1. Recall that the estimation codes methods often used for fitting models for computer experiments from deterministic codes have the characteristic of no uncertainty at an input location and increasing uncertainty as we move farther away from any observation. Hence, space-filling designs are desirable since in general they spread the design points out nearly evenly or uniformly (in some sense) throughout the region of experimentation. This is also a desirable feature if the experimenter doesn't know the form of the model that is required and believes that interesting phenomena are likely to be found throughout all regions of the experimental space.
2. Furthermore, most space-filling designs do not contain any replicate runs, even when the design is projected into lower dimensional spaces. For a deterministic computer model this is desirable, because a single run of the computer model at a design point provides all of the information about the response at that point. If the design were to contain replicates when projected into lower dimensions and some of the factors were not active, then this could result in the same response value being obtained for multiple runs. Since it is not known a priori which factors are active, this could be costly duplication if obtaining those runs requires a great deal of computational effort.
3. Some regions of the design space may be known to be inadmissible based on underlying science or engineering knowledge. Space-filling designs can be easily adapted to fit into nonstandard shaped regions.



**Figure 10.5** A 10-run Latin hypercube design.

The first space-filling design proposed was the Latin hypercube design [McKay, Conover, and Beckman (1979)]. A Latin hypercube (LHC) in  $n$  runs for  $k$  factors is an  $n \times k$  matrix where each column is a permutation of the levels  $1, 2, \dots, n$ , which is then scaled and centered to lie in the desired interval. JMP can create Latin hypercube designs. An example of a 10-run Latin hypercube design in two factors from JMP on the interval  $-1$  to  $+1$  is shown in Fig. 10.5. The 10 chosen levels of the design are  $-1, -7/9, -5/9, -3/9, -1/9, \dots, 7/9, 1$ . An arbitrary permutation of these levels for each factor might yield an undesirable filling of the space. Consider the worst-case LHC design with design locations  $(-1, -1), (-7/9, -7/9), \dots, (1, 1)$ . This design has all locations lying along a single line and has no data points in the remainder of the design space. Hence a secondary criterion is used to ensure that the design selected does indeed have good space-filling properties. Two common choices for ensuring good space-filling for an LHC (or other possible design construction strategies) are:

- **Minimax:** A minimax design is one that minimizes the maximum distance between any location in the design space and its nearest design point. This criterion directly targets the objective of allowing a design point to never be too far away from a new location where we may wish to predict.
- **Maximin:** A maximin design is one that maximizes the minimum distance between any two design points. This criterion forces neighboring points to be as far apart as possible, thereby filling the space. This criterion has the advantage of allowing comparison of distances to just involve the design points, instead of all possible locations in the design space.

The two criteria often produce similar results, with the maximin criterion being more popular with its ease of calculation and implementation. The design shown in Fig. 10.5 is a LHC design that achieves its optimal spacing using the maximin criterion.

A variation of the Latin hypercube design described in Santner, Williams, and Notz (2003) allows additional flexibility in the choice of the levels of each factor selected. Instead of using the fixed locations of  $1, 2, \dots, n$  for an  $n$  run design. It partitions the factor range for each input into  $n$  equal width intervals  $[0, 1/n), [1/n, 2/n), \dots, [(n-1)/n, 1]$ . A good

LHC is constructed using the maximin or minimax criterion, and to define which regions in which to place a point, and then the particular value of each factor is selected from a uniform distribution within each range.

Other possible space filling designs include:

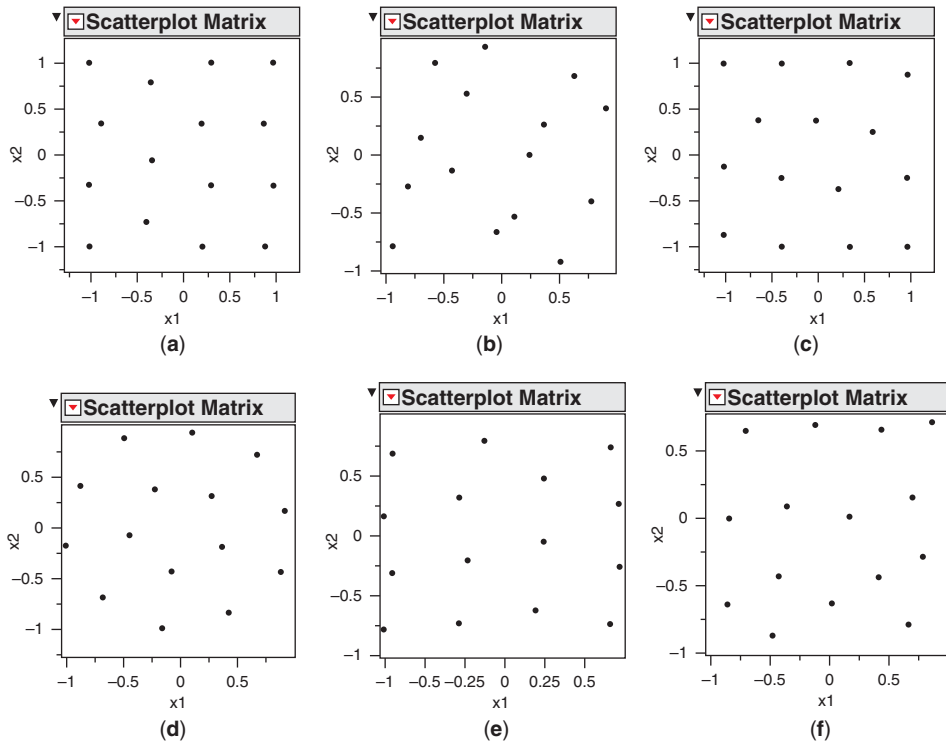
1. Sphere-packing designs proposed by Johnson et al. (1990). These designs use the maximin criterion described above, but look across the set of all possible design locations, instead of being restricted to just those with the Latin hypercube structure.
2. Uniform designs were proposed by Fang (1980) and attempt to place the design points so that they are uniformly scattered through the region as would a sample from a uniform distribution. For a number of algorithms for creating these designs and several measures of uniformity, see the book by Fang et al. (2006).
3. Maximum entropy designs were proposed by Shewry and Wynn (1987). Entropy can be thought of as a measure of the amount of information contained in the distribution of a data set. Suppose that the data comes from a normal distribution with mean vector  $\mu$  and covariance matrix  $\sigma^2\mathbf{R}(\theta)$ , where  $\mathbf{R}(\theta)$  is a correlation matrix having elements

$$r_{ij} = e^{-\sum_{s=1}^k \theta_s(x_{is}-x_{js})^2} \quad (10.30)$$

The quantities  $r_{ij}$  are the correlations between the responses at two design points. The maximum entropy design maximizes the determinant of  $\mathbf{R}(\theta)$ .

4. Minimum potential designs are also included in JMP that have properties that can be described based on a physics analogy. We consider the design points as electrons jointed by springs. Their common charge forces them apart, while the springs try to pull them together. The minimum potential design selects those points that minimize the potential energy of the system.
5. Gaussian process IMSE designs were described by Welch et al. (1992) and Jones and Johnson (2009). These designs minimize the integrate mean square error for a standard Gaussian process model (described in the next section) over the design space.
6. Fast Flexible Filling (FFF) designs are also included in JMP. These designs are highly adaptable for nonstandard design spaces with constraints or inadmissible regions.

In lower dimensions, several of the above choices produce similar designs. Figure 10.6 shows the six alternatives for a 15-run design in two factors from JMP. As the number of input factors increases, the distinctions between the alternatives grow. Figure 10.7 shows the pairwise scatterplots for Latin hypercube and Fast flexible filling designs for a 40-run design in six factors from JMP. Not how the LHC design continues to project well into lower dimensions, while the FFF places more emphasis on placing points closer to the edges of the design space, since in higher dimensions there is more volume there. Johnson, Montgomery, and Jones (2011) compare the performance of different choices of computer designs across different scenarios.



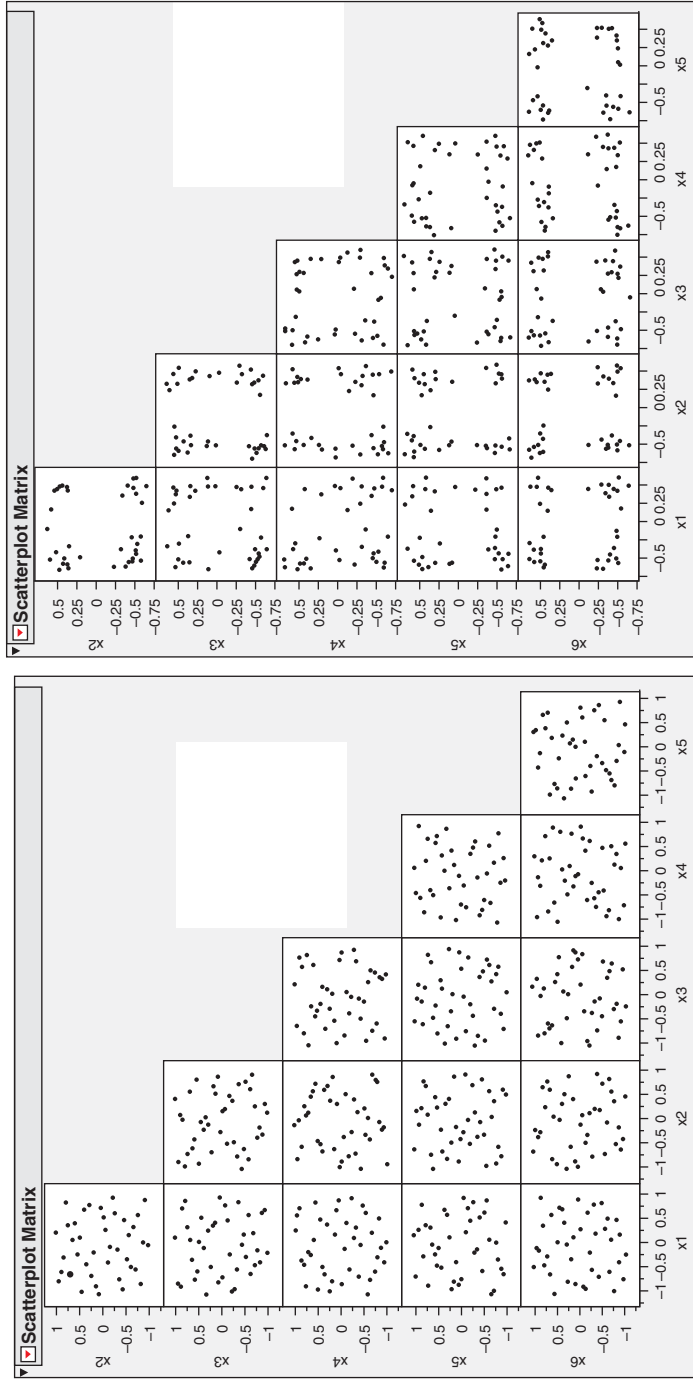
**Figure 10.6** Alternative designs for a 15-run design with two factors: (a) Sphere packing. (b) Uniform. (c) Maximum entropy. (d) Minimum potential. (e) Gaussian process IMSE optimal. (f) Fast flexible filling.

#### 10.4.2 Analysis for Computer Experiments

The Gaussian process model is often used to fit the data from a deterministic computer experiment. These models were introduced as models for computer experiments by Sacks et al. (1989a). They are desirable because they provide an exact fit to the observations from the experiment. Now this is no assurance that they will interpolate well at locations in the region of interest where there is no data, and no one seriously believes that the Gaussian process model is the correct model for the relationship between the response and the design variables. However, the “exact fit” nature of the model and the fact that it only requires one parameter for each factor considered in the experiment has made it quite popular. The Gaussian process model is

$$y = \mu + z(\mathbf{x})$$

where  $z(\mathbf{x})$  is a Gaussian stochastic process with covariance matrix  $\sigma^2 \mathbf{R}(\boldsymbol{\theta})$ , and the elements of  $\mathbf{R}(\boldsymbol{\theta})$  are defined in Equation 10.30. The Gaussian process model is essentially a spatial correlation model, where the correlation of the response between two observations decreases as the values of the design factors become further apart. When design points are close together, this causes ill-conditioning in the data for the Gaussian process model, much like multicollinearity resulting from predictors that are nearly linearly dependent in linear



**Figure 10.7** Pairwise scatter plots for alternative designs for a 40-run design with six factors: (a) Latin hypercube. (b) Fast flexible filling.

**TABLE 10.3** Sphere-Packing Design and the Temperature Responses in the CFD Experiment

| <i>x</i> -axis | <i>R</i> -axis | Temperature |
|----------------|----------------|-------------|
| 0.056          | 0.062          | 338.07      |
| 0.095          | 0.013          | 1613.04     |
| 0.077          | 0.062          | 335.91      |
| 0.095          | 0.061          | 327.82      |
| 0.090          | 0.037          | 449.23      |
| 0.072          | 0.038          | 440.58      |
| 0.064          | 0.015          | 1173.82     |
| 0.050          | 0.000          | 1140.36     |
| 0.050          | 0.035          | 453.83      |
| 0.079          | 0.000          | 1261.39     |

regression models. The parameters  $\mu$  and  $\theta_s$ ,  $s = 1, 2, \dots, k$  are estimated using the method of maximum likelihood. Predicted values of the response at the point  $\mathbf{x}$  are computed from

$$\hat{y}(\mathbf{x}) = \hat{\mu} + \mathbf{r}'(\mathbf{x})\mathbf{R}(\hat{\theta})^{-1}(\mathbf{y} - \mathbf{j}\hat{\mu}) \quad (10.31)$$

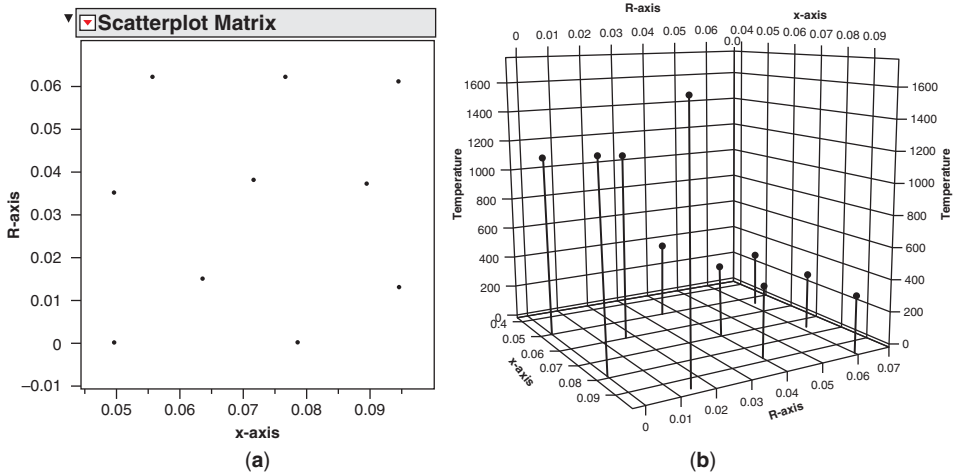
where  $\hat{\mu}$  and  $\hat{\theta}$  are the maximum likelihood estimates of the model parameters  $\mu$  and  $\theta$ , and  $\mathbf{r}'(\mathbf{x}) = [\mathbf{r}(\mathbf{x}_1, \mathbf{x}), \mathbf{r}(\mathbf{x}_2, \mathbf{x}), \dots, \mathbf{r}(\mathbf{x}_n, \mathbf{x})]$ . The prediction equation contains one model term for each design point in the original experiment. JMP will fit and provide predictions from the Gaussian process model.

**Example 10.2** The temperature in the exhaust from a jet turbine engine at different locations in the plume was studied using a computational fluid dynamics (CFD) model. The two design factors of interest were the locations in the plume ( $x$  and  $y$  coordinates, however the  $y$  axis was referred to by the experimenters as the  $R$ -axis or radial axis). Both location axes were coded to the  $-1, +1$  interval. The experimenters used a 10-run sphere-packing design. The experimental design and the output obtained at these test conditions from the CFD model are shown in Table 10.3. Figure 10.8 shows the design and response values.

JMP was used to fit the Gaussian process model to the temperature data. Some of the output is shown in Table 10.4 and Fig. 10.9. The surface plot in Fig. 10.9a shows the overall surface throughout the entire design space. The plot of actual by predicted in Fig. 10.9b is obtained by “jack-knifing” the predicted values; that is, each predicted value is obtained from a model that doesn’t contain that observation when the model parameters are estimated. The marginal model plots of Fig. 10.9c show that the predicted values from the model change much more as we vary the value of the  $R$ -axis compared to the  $x$ -axis.

The prediction model obtained from JMP has a symmetric structure with a term of the same form for each observed value in the design is shown below, with “ $x$ ” and “ $R$ ” denoting the “ $x$ -axis” and “ $R$ -axis” from Table 10.3.

$$\begin{aligned} \hat{y} = & 736.3228 \\ & -2090.5005 * \exp(-(65.7321 * (x - 0.056)^2 + 3360.7699 * (R - 0.062)^2)) \\ & + 3944.7842 * \exp(-(65.7321 * (x - 0.095)^2 + 3360.7699 * (R - 0.013)^2)) \end{aligned}$$



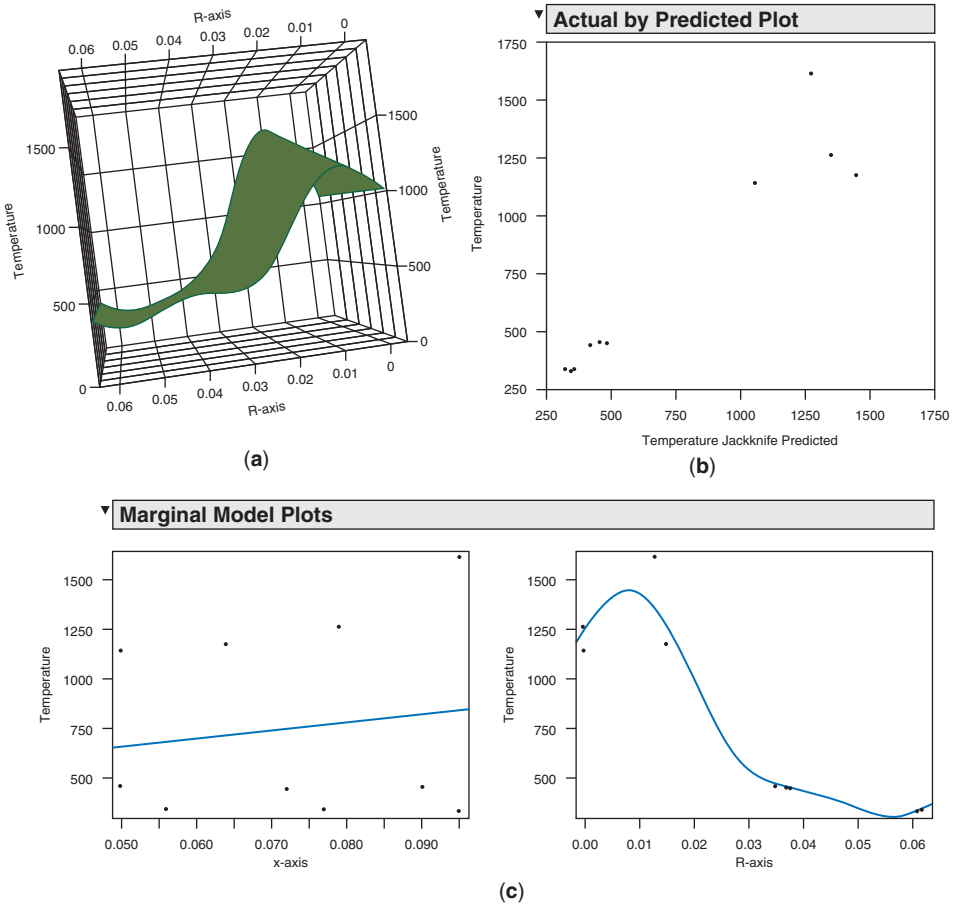
**Figure 10.8** (a) The sphere-packing design for the CFD experiment, and (b) a 3-D plot of the observed responses.

$$\begin{aligned}
 &+3806.0602 * \exp(-(65.7321 * (x - 0.077)^2 + 3360.7699 * (R - 0.062)^2)) \\
 &-2235.3120 * \exp(-(65.7321 * (x - 0.095)^2 + 3360.7699 * (R - 0.061)^2)) \\
 &-1065.8557 * \exp(-(65.7321 * (x - 0.09)^2 + 3360.7699 * (R - 0.037)^2)) \\
 &+975.0887 * \exp(-(65.7321 * (x - 0.072)^2 + 3360.7699 * (R - 0.038)^2)) \\
 &-3074.0849 * \exp(-(65.7321 * (x - 0.064)^2 + 3360.7699 * (R - 0.015)^2)) \\
 &+1041.2588 * \exp(-(65.7321 * (x - 0.05)^2 + 3360.7699 * R^2)) \\
 &-69.2146 * \exp(-(65.7321 * (x - 0.05)^2 + 3360.7699 * (R - 0.035)^2)) \\
 &-1232.2241 * \exp(-(65.7321 * (x - 0.079)^2 + 3360.7699 * R^2))
 \end{aligned}$$

Other useful references for designs for computer experiments include Bettonvil and Kleijnen (1996), Donohue (1994), Montgomery and Weatherby (1979), Sacks et al. (1989a and b), Sacks and Welch (1989), Slaagame and Barton (1997), and the book by Santner et al. (2003).

**TABLE 10.4 JMP Output for the Gaussian Process Model for the CFD Experiment**

| Model Report                                 |            |                   |             |                    |                    |
|--|------------|-------------------|-------------|--------------------|--------------------|
| Column                                       | Theta      | Total Sensitivity | Main Effect | x-axis Interaction | R-axis Interaction |
| x-axis                                       | 65.732136  | 0.0370478         | 0.0147994   |                    | 0.0222485          |
| R-axis                                       | 3360.7699  | 0.9852006         | 0.9629522   | 0.0222485          |                    |
| $\mu$  | $\sigma^2$ |                   |             |                    |                    |
| 736.3228                                     | 214612.38  |                   |             |                    |                    |
| -2*LogLikelihood                             | 132.9562   |                   |             |                    |                    |
| Fit using the Gaussian correlation function. |            |                   |             |                    |                    |



**Figure 10.9** JMP output for the Gaussian process model for the CFD experiment in Table 10.3. (a) Surface plot of the predicted response. (b) Actual by predicted plot. (c) Marginal model plots for *x*-axis and *R*-axis.

### 10.4.3 Combining Information from Physical and Computer Experiments

As described at the beginning of Section 10.4, the characteristics of data from physical experiments and computer codes are very different. In some applications, there can be some advantages to combining data from both these sources to obtain results that leverage the advantages of both types of data. For example, if some data from physical experiments are available, then this can provide direct information about how the system actually performs. However, it may be prohibitively expensive to obtain large amounts of this type of information, particular in some parts of the region of operability. In addition, these data will have some natural variability associated with them. Alternately, obtaining data from computer codes can be obtained in different regions of interest and can often be less expensive to obtain if the code is already available. The weakness of this data is that, depending on how mature the computer code is that generates the data, there may be some bias from what the true underlying mechanism is.



By combining information from both sources of data, it may be possible to obtain improved predictions for new observations as well as provide feedback to the scientists and engineers who developed the computer code, to allow them to improve the model and make it a closer match to reality. A common form of the model is

$$Y = f(x; \theta) + \delta(x) + \varepsilon$$

where  $f(x; \theta)$  is the result from the computer code at an input combination,  $x$ , and  $\varepsilon$  is the error term associated with the physical data. The discrepancy term,  $\delta(x)$ , is used to describe systematic differences between the observed physical data and what is estimated from the computer code.

The discrepancy term both corrects for systematic bias from the computer model, which might result from incomplete knowledge of the underlying process, and provides a summary that can be studied by the subject matter experts to suggest improvements or updates to the model. Joseph and Yan (2014) describe a process for continuous improvement by exploring the shape of the discrepancy to help guide the next generation of computer code.

Useful references on this topic include Kennedy and O'Hagan (2001), Higdon et al. (2004, 2008), Williams et al. (2006), and Weaver et al. (2015).

## 10.5 MINIMUM BIAS ESTIMATION OF RESPONSE SURFACE MODELS

In Section 10.2 we considered a situation where the experimenter fits a model, say

$$\hat{y} = \mathbf{X}_1 \mathbf{b} \quad (10.32)$$

that is of lower order than the true system model

$$\mathbf{f}(\mathbf{x}) = E(\mathbf{y}) = \mathbf{X}_1 \boldsymbol{\beta}_1 + \mathbf{X}_2 \boldsymbol{\beta}_2 \quad (10.33)$$

This leads to the AMSE criterion in Equation 10.11, which can be written as

$$\text{AMSE} = \text{APV} + \text{ASB} \quad (10.34)$$

where APV is the average prediction variance and ASB is the average squared bias. We discussed how to minimize the ASB portion of AMSE by properly selecting the experimental design. We also discussed the important role that the bias plays in the AMSE criterion, noting that often a design that is close to the minimum bias design comes very close to the design that minimizes the AMSE.

Karson, Manson, and Hader (1969) proposed minimizing the average squared bias not by design, but by **estimation**. Their **minimum bias estimation** procedure achieves the minimum value of ASB, leaving the experimenter to select the design to satisfy other objectives such as the minimization of average prediction variance. We now describe their procedure. The true model is a polynomial of degree  $d_2$ ,

$$f(\mathbf{x}) = \mathbf{x}'_1 \boldsymbol{\beta}_1 + \mathbf{x}'_2 \boldsymbol{\beta}_2 \quad (10.35)$$

and the fitted model is a polynomial of degree  $d_1 < d_2$ ,

$$\hat{y}(\mathbf{x}) = \mathbf{x}'_1 \mathbf{b}_1 \quad (10.36)$$

Recall from Equation 10.11 that the average squared bias is

$$\text{ASB} = \frac{NK}{\sigma^2} \int_R \{E[\hat{y}(\mathbf{x}) - f(\mathbf{x})]\}^2 dx \quad (10.37)$$

Now  $E[\hat{y}(\mathbf{x})]$  is a polynomial of degree  $d_1$ , which we may write as

$$E[\hat{y}(\mathbf{x})] = \mathbf{x}'_1 \boldsymbol{\alpha}_1$$

Consequently, the average squared bias is

$$\begin{aligned} \text{ASB} &= \frac{NK}{\sigma^2} \int_R \{E[\hat{y}(\mathbf{x}) - f(\mathbf{x})]\}^2 d\mathbf{x} \\ &= \frac{NK}{\sigma^2} \int_R [\mathbf{x}'_1(\boldsymbol{\alpha}_1 - \boldsymbol{\beta}_1) - \mathbf{x}'_2 \boldsymbol{\beta}_2]^2 d\mathbf{x} \\ &= \frac{NK}{\sigma^2} \int_R \{(\boldsymbol{\alpha}_1 - \boldsymbol{\beta}_1)' \mathbf{x}_1 \mathbf{x}'_1 (\boldsymbol{\alpha}_1 - \boldsymbol{\beta}_1) - 2(\boldsymbol{\alpha}_1 - \boldsymbol{\beta}_1) \mathbf{x}_1 \mathbf{x}'_2 \boldsymbol{\beta}_2 + \boldsymbol{\beta}'_2 \mathbf{x}_2 \mathbf{x}'_2 \boldsymbol{\beta}_2\} d\mathbf{x} \\ &= \frac{N}{\sigma^2} [(\boldsymbol{\alpha}_1 - \boldsymbol{\beta}_1)' \boldsymbol{\mu}_{11} - 2(\boldsymbol{\alpha}_1 - \boldsymbol{\beta}_1) \boldsymbol{\mu}_{12} \boldsymbol{\beta}_2 + \boldsymbol{\beta}'_2 \boldsymbol{\mu}_{22} \boldsymbol{\beta}_2] \end{aligned} \quad (10.38)$$

where  $\boldsymbol{\mu}_{11}$ ,  $\boldsymbol{\mu}_{12}$ , and  $\boldsymbol{\mu}_{22}$  are the region moment matrices defined earlier. Now to minimize the ASB with respect to  $\boldsymbol{\alpha}_1$ , differentiate Equation 10.38 with respect to  $\boldsymbol{\alpha}_1$ , equate the result to  $\mathbf{0}$ , and solve:

$$\frac{\partial \text{ASB}}{\partial \boldsymbol{\alpha}_1} = 2\boldsymbol{\mu}_{11}(\boldsymbol{\alpha}_1 - \boldsymbol{\beta}_1) - 2\boldsymbol{\mu}_{12}\boldsymbol{\beta}_2 = \mathbf{0}$$

which results in

$$\begin{aligned} \boldsymbol{\alpha}_1 &= \boldsymbol{\beta}_1 + \boldsymbol{\mu}_{11}^{-1} \boldsymbol{\mu}_{12} \boldsymbol{\beta}_2 \\ &= \mathbf{A} \boldsymbol{\beta} \end{aligned} \quad (10.39)$$

where  $\mathbf{A} = [\mathbf{I} \mid \boldsymbol{\mu}_{11}^{-1} \boldsymbol{\mu}_{12}]$  and  $[\boldsymbol{\beta}'_1 \mid \boldsymbol{\beta}'_2]$ . So a necessary and sufficient condition for minimization of average squared bias is that

$$E[\hat{y}(\mathbf{x})] = \mathbf{x}'_1 \mathbf{A} \boldsymbol{\beta} \quad (10.40)$$

or

$$E(\mathbf{b}_1) = \mathbf{A} \boldsymbol{\beta} \quad (10.41)$$

Now  $\mathbf{b}_1$  is just a linear combination of the observations, say

$$\mathbf{b}_1 = \mathbf{T}' \mathbf{y} \quad (10.42)$$

Because  $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$ , the requirement that  $E(\mathbf{b}_1) = \mathbf{A}\boldsymbol{\beta}$  means that  $\mathbf{T}'$  must satisfy

$$\mathbf{T}'\mathbf{X} = \mathbf{A} \quad (10.43)$$

where  $\mathbf{X} = [\mathbf{X}_1 \ : \ \mathbf{X}_2]$ . The minimum value of the average squared bias is

$$\min \text{ASB} = \frac{N}{\sigma^2} \boldsymbol{\beta}'_2 [\boldsymbol{\mu}_{22} - \boldsymbol{\mu}'_{12} \boldsymbol{\mu}_{11}^{-1} \boldsymbol{\mu}_{12}] \boldsymbol{\beta}_2. \quad (10.44)$$

In addition to achieving the minimum value of ASB, we also want to minimize the average prediction variance (APV) in Equation 10.11. Now the minimum variance estimator of  $\mathbf{x}'_1 \mathbf{A}\boldsymbol{\beta}$  is

$$\hat{y}(\mathbf{x}) = \mathbf{x}'_1 \mathbf{A}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} \quad (10.45)$$

The APV is minimized when Equation 10.45 is satisfied, or when

$$\mathbf{T}' = \mathbf{A}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \quad (10.46)$$

The variance of the fitted model is

$$\text{Var}[\hat{y}(\mathbf{x})] = \sigma^2 \mathbf{x}'_1 \mathbf{T}' \mathbf{T} \mathbf{x}_1 \quad (10.47)$$

**Example 10.3 One Design Variable** Suppose we have one design variable and we wish to fit a first-order model using three observations taken at the points  $x = 0$  and  $x = \pm 1$ . The region of interest is  $R = [-1, 1]$ , and, unknown to us, the true model is quadratic. Therefore,

$$\begin{aligned} \hat{y}(\mathbf{x}) &= \mathbf{x}'_1 \mathbf{b}_1 = b_0 + b_1 x \\ f(\mathbf{x}) &= \mathbf{x}'_1 \boldsymbol{\beta}_1 + \mathbf{x}'_2 \boldsymbol{\beta}_2 = (\beta_0 + \beta_1 x) + \beta_2 x^2 \end{aligned}$$

with  $\mathbf{x}'_1 = [1, x]$ ,  $\mathbf{x}'_2 = [x^2]$ ,  $\boldsymbol{\beta}'_1 = [\beta_0, \beta_1]$ ,  $\boldsymbol{\beta}'_2 = [\beta_2]$ , and  $\mathbf{b}'_1 = [b_0, b_1]$ . Also, note that

$$\begin{aligned} \mathbf{X} = [\mathbf{X}_1 \ : \ \mathbf{X}_2] &= \begin{bmatrix} 1 & -1 & \vdots & 1^2 \\ 1 & 0 & \vdots & 0 \\ 1 & 1 & \vdots & 1^2 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \\ (\mathbf{X}'\mathbf{X})^{-1} &= \begin{bmatrix} 1 & 0 & -l^{-2} \\ 0 & \frac{1}{2}l^{-2} & 0 \\ -l^{-2} & 0 & \frac{3}{2}l^{-4} \end{bmatrix} \\ k = \frac{1}{2}, \quad \boldsymbol{\mu}_{11} &= \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{bmatrix}, \quad \boldsymbol{\mu}_{12} = \begin{bmatrix} \frac{1}{3} \\ 0 \end{bmatrix}, \quad \boldsymbol{\mu}_{22} = \begin{bmatrix} \frac{1}{5} \end{bmatrix} \end{aligned}$$

and

$$\mathbf{A} = [\mathbf{I}_2 \ : \ \boldsymbol{\mu}_{11}^{-1} \boldsymbol{\mu}_{12}] = \begin{bmatrix} 1 & 0 & \vdots & \frac{1}{3} \\ 0 & 1 & \vdots & 0 \end{bmatrix}$$

Therefore the minimum bias estimate of  $\mathbf{b}_1$  is

$$\begin{aligned}\mathbf{b}_1 &= \mathbf{T}'\mathbf{y} \\ &= \mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ &= \frac{1}{6l^2} \begin{bmatrix} 1 & 2(3l^2 - 1) & 1 \\ -3l & 0 & 3l \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}\end{aligned}$$

The minimum bias estimator is biased, since

$$E(\mathbf{b}_1) = \mathbf{A}\boldsymbol{\beta} = \begin{bmatrix} \beta_0 + \frac{1}{3}\beta_2 \\ \beta_1 \end{bmatrix}$$

and the predicted values are also biased, because

$$E[\hat{y}(\mathbf{x})] = \beta_0 + \frac{1}{3}\beta_2 + \beta_1 x$$

The minimum value of the ASB is computed from Equation 10.44 as

$$\begin{aligned}\text{Min ASB} &= \frac{N}{\sigma^2} \boldsymbol{\beta}'_2 [\boldsymbol{\mu}_{22} - \boldsymbol{\mu}'_{12} \boldsymbol{\mu}_{11}^{-1} \boldsymbol{\mu}_{12}] \boldsymbol{\beta}_2 \\ &= \left(\frac{3}{\sigma^2}\right) \frac{4}{45} \beta_2^2\end{aligned}$$

Notice that the minimum bias is independent of where the design points  $-l, +l$  are placed in the region  $-1 \leq x \leq +1$ .

The variance of the fitted model is

$$\begin{aligned}\text{Var}[\hat{y}(\mathbf{x})] &= \sigma^2 \mathbf{x}'_1 \mathbf{T}' \mathbf{T} \mathbf{x}_1 \\ &= \sigma^2 \mathbf{x}'_1 \mathbf{A}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{A} \mathbf{x}_1\end{aligned}$$

and we can show that the APV is

$$\text{APV} = \frac{Nk}{\sigma^2} \int_{-1}^1 \text{Var}[\hat{y}(\mathbf{x})] d\mathbf{x} = 3 - \frac{3}{2l^2} + \frac{1}{2l^4}$$

The APV is minimized at  $l = \sqrt{\frac{2}{3}}$ , for which value the minimum APV is  $\frac{15}{8}$ .

We can compare this with the usual least squares results. It is not difficult to show that

$$\text{ASB}_{\text{LS}} = \frac{3\beta_2^2}{\sigma^2} \left( \frac{4l^4}{9} - \frac{4l^2}{9} + \frac{1}{5} \right)$$

and

$$\text{APV}_{\text{LS}} = 1 + \frac{1}{2l^2}$$

Now it turns out that  $APL_{LS}$  is less than  $APV$  for minimum bias estimation except at  $l = \sqrt{\frac{1}{2}}$ , where they are identical. The minimum  $ASB_{LS}$  is achieved at  $l = \sqrt{\frac{2}{3}}$ , while the  $ASB$  is minimized for *any*  $l$  for minimum bias estimation. However, for any value of  $l$  satisfying  $\sqrt{\frac{1}{2}} \leq l \leq 1$  we have  $APV \leq APV_{LS} = 2$ . Therefore, if we use minimum bias estimation, the average mean squared error is less than or equal to the average mean squared error for least squares for *any* design with  $\sqrt{\frac{1}{2}} \leq l \leq 1$ .

## 10.6 NEURAL NETWORKS

Some researchers have suggested using **neural networks** as an alternative to RSM; for example, see Han et al. (1994), Himmel and May (1991), Carpenter and Barthelemy (1993), and Hussain, Yu, and Johnson (1991). The development of neural networks, or more accurately, **artificial neural networks**, has been motivated by the recognition that the human brain processes information in a way that is fundamentally different from the typical digital computer. The neuron is the basic structural element and information-processing module of the brain. A typical human brain has an enormous number of them (approximately 10 billion neurons in the cortex and 60 trillion synapses or connections between them) arranged in a highly complex, nonlinear, and parallel structure. Consequently, the human brain is a very efficient structure for information processing, learning, and reasoning.

An artificial neural network is a structure that is designed to solve certain types of problems by attempting to emulate the way the human brain would solve the problem. The general form of a neural network is a “black box” model of a type that is often used to model high-dimensional, nonlinear data. Typically, most neural networks are used to solve prediction problems for some system, as opposed to formal model-building or development of underlying knowledge of how the system works. For example, a computer company might want to develop a procedure for automatically reading handwriting and converting it to typescript. If the procedure can do this quickly and accurately, the company may have little interest in understanding the functional form of the specific model used to do it.

Multilayer feedforward artificial neural networks are multivariate statistical models used to relate  $p$  predictor variables  $x_1, x_2, \dots, x_p$  to  $q$  response variables  $y_1, y_2, \dots, y_q$ . The model has several layers, each consisting of either the original or some constructed variables. The most common structure involves three layers: the **inputs**, which are the original predictors; the **hidden layer**, consisting of a set of constructed variables; and the **output layer**, made up of the responses. Each variable in a layer is called a **node**. Figure 10.10 shows a typical three-layer artificial neural network with inputs on the left and outputs on the right.

A node takes as its input a transformed linear combination of the outputs from the nodes in the previous layer. Then it sends as an output a transformation of itself that becomes one of the inputs to one or more nodes in the next layer. The transformation functions are usually either sigmoidal (S-shaped) or linear, and are usually called **activation functions** or **transfer functions**. Let each of the  $k$  hidden layer nodes  $a_u$  be a linear combination of the input variables:

$$a_u = \sum_{j=1}^p w_{1ju} x_j + \theta_u$$

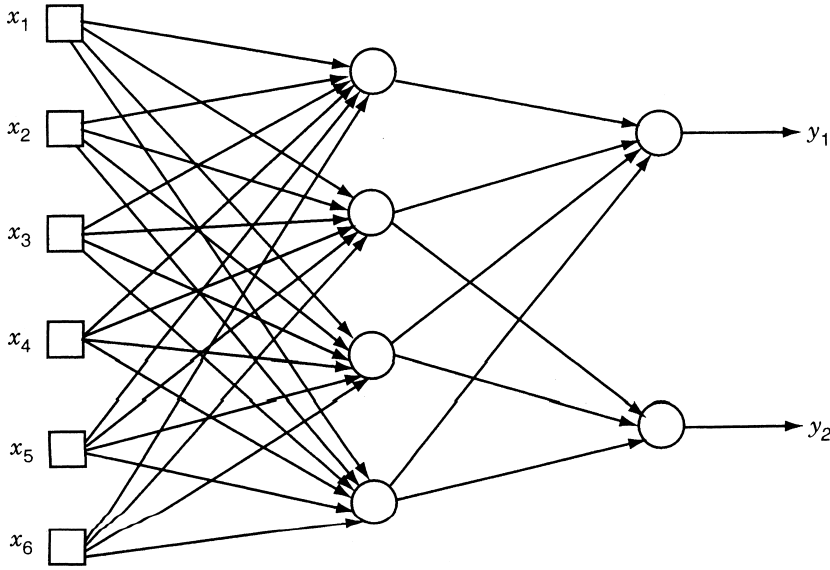


Figure 10.10 Artificial neural network with one hidden layer.

where the  $w_{1ju}$  are unknown parameters that must be estimated (called **weights**) and  $\theta_u$  is a parameter that plays the role of an intercept in linear regression (this parameter is sometimes called the **bias node**).

Each node is transformed by the activation function  $g(\cdot)$ . Much of the neural networks literature refers to these activation functions notationally as  $\sigma(\cdot)$  because of their S-shape (this is an unfortunate choice of notation so far as statisticians are concerned).

Let the output of node  $a_u$  be denoted by  $z_u = g(a_u)$ . Now we form a linear combination of these outputs, say  $b_v = \sum_{u=0}^k w_{2uv}z_u$ , where  $z_0 = 1$ . Finally, the  $v$ th response  $y_v$  is a transformation of the  $b_v$ , say  $y_v = \tilde{g}(b_v)$ , where  $\tilde{g}(\cdot)$  is the activation function for the response. This can all be combined to give

$$y_v = \tilde{g} \left[ \sum_{k=1}^u w_{2uv}g \left( \sum_{j=1}^p w_{1ju}x_j + \theta_{1j} \right) + \theta_{2u} \right] \tag{10.48}$$

The response  $y_v$  is a transformed linear combination of transformed linear combinations of the original predictors. For the hidden layer, the activation function is often chosen to be either the logistic function  $g(x) = 1/(1 + e^{-x})$  or the hyperbolic tangent function  $g(x) = \tanh(x) = (e^x - e^{-x})/(e^x + e^{-x})$ . The choice of activation function for the output layer depends on the nature of the response. If the response is bounded or dichotomous the output activation function is usually taken to be sigmoidal, while if it is continuous an identity function is often used.

The model in Equation 10.48 is a very flexible form containing many parameters, and it is this feature that gives a neural network a nearly universal approximation property. That is, it will fit many naturally occurring functions. However, the parameters in 10.48 must be estimated, and there are a lot of them. The usual approach is to estimate the parameters by minimizing the overall residual sum of squares taken over all responses and all observations. This is a nonlinear least squares problem, and a variety of algorithms can be used to solve

it. Often a procedure called **backpropagation** (which is a variation of steepest descent) is used, although derivative-based gradient methods have also been employed. As in any nonlinear estimation procedure, starting values for the parameters must be specified in order to use these algorithms. It is customary to standardize the ranges of all the input variables so that small, essentially random values are chosen for the starting values.

With so many parameters involved in a complex nonlinear function, there is considerable danger of **overfitting**. That is, a neural network will provide a nearly perfect fit to a set of historical or **training** data, but it will often predict new data very poorly. Overfitting is a familiar problem to statisticians trained in empirical model building. The neural network community has developed various methods for dealing with this problem, such as reducing the number of unknown parameters (this is called **optimal brain surgery**), stopping the parameter estimation process before complete convergence and using cross-validation to determine the number of iterations to use, and adding a penalty function to the residual sum of squares that increases as a function of the sum of the squares of the parameter estimates. There are also many different strategies for choosing the number of layers and number of neurons and the form of the activation functions. This is usually referred to as choosing the **network architecture**. Cross-validation can be used to select the number of nodes in the hidden layer. Good references on artificial neural networks are Cheng and Titterton (1994), Bishop (1995), Haykin (1994), and Ripley (1994a, 1994b).

Artificial neural networks are an active area of research and application, particularly for the analysis of large, complex, highly nonlinear problems. The overfitting issue is frequently overlooked by many users and advocates of neural networks. Because many members of the neural network community do not have sound training in empirical model building, they often do not appreciate the difficulties overfitting may cause. Furthermore, many computer programs for implementing neural networks do not handle the overfitting problem particularly well. Our view is that neural networks can be a complement to the familiar statistical tools of regression analysis, RSM, and designed experiments, but certainly not a replacement for them. A neural network can at best only give a prediction model but not fundamental insight into the underlying process mechanism that produced the data. Furthermore, there is no reason to believe that the prediction properties of these models are superior to those that would be obtained from a well-designed RSM study.

## 10.7 SPLIT-PLOT DESIGNS FOR SECOND-ORDER MODELS

In Chapters 3 and 4 we introduced the split-plot design (SPD) as a technique for running experiments more economically if some of the factors are hard to change. Time or cost constraints lead to considering a split-plot structure rather than a completely randomized design. For response surface models involving second-order models, practical constraints may often dictate that a split-plot design and model be considered to accommodate some hard-to-change factors. Recall that SPDs involve two randomizations [Letsinger et al. (1996) and Lucas and Ju (1992)]. The hard-to-change factor combinations are randomly assigned to whole-plot units based on the whole plot design. Within each whole plot, the easy-to-change factors are assigned to the subplot units with a separate randomization for each whole plot. These two levels of randomization lead to two error terms, a whole plot error and a subplot error. Care should be taken to ensure that both types of error can be estimated, and that the terms of the model are assessed relative to the appropriate error term. Typically, the main, interaction, and quadratic effects involving whole plot factors

**TABLE 10.5** A Restricted Central Composite Split-Plot Design with Nine Whole Plots for Two Hard-to-Change and Two Easy-to-Change Factors

| Whole Plot # | $z_1$     | $z_2$     | $x_1$       | $x_2$       | # Runs  |
|--------------|-----------|-----------|-------------|-------------|---------|
| 1            | -1        | -1        | $\pm 1$     | $\pm 1$     | 4       |
| 2            | -1        | +1        | $\pm 1$     | $\pm 1$     | 4       |
| 3            | +1        | -1        | $\pm 1$     | $\pm 1$     | 4       |
| 4            | +1        | +1        | $\pm 1$     | $\pm 1$     | 4       |
| 5            | $-\alpha$ | 0         | 0           | 0           | 1       |
| 6            | $+\alpha$ | 0         | 0           | 0           | 1       |
| 7            | 0         | $-\alpha$ | 0           | 0           | 1       |
| 8            | 0         | $+\alpha$ | 0           | 0           | 1       |
| 9            | 0         | 0         | $\pm\alpha$ | 0           | 4 + #CR |
|              |           |           | 0           | $\pm\alpha$ |         |
|              |           |           | <b>0</b>    | <b>0</b>    |         |

are whole-plot effects. Terms of the model involving just subplot factors or combinations of whole and subplot factors are generally subplot effects.

Split-plot designs have practical advantages for running the experiment since the hard-to-change factors are only reset for each whole plot, while the easy-to-change factors are reset for each observation. Not only are these designs more economical to run, but as Goos and Vanderbroek (2004) show, they can also sometimes be more  $D$ - and  $G$ -efficient than a similar sized completely randomized design.

The central composite and Box–Behnken designs presented in Chapter 8 can be easily adapted to become split-plot designs with good properties. One simple approach to adapting these designs is to group all experimental runs with common whole-plot levels into a single whole plot. These **restricted split-plot designs** have the minimal allowable number of whole plots, with just a single whole plot for each combination of whole plot levels. Table 10.5 shows an example of a restricted central composite design for two hard-to-change ( $z_1$  and  $z_2$ ) and two easy-to-change ( $x_1$  and  $x_2$ ) factors. Table 10.6 shows an example of a restricted Box–Behnken design for the same set of factors. If the hard-to-change factors are prohibitively expensive, then this approach allows us to minimize the cost of changing these factors. One obvious weakness of this type of design is that there are no replicated whole plots to estimate the whole-plot error term.

In Section 8.1, some important characteristics for a good response surface design are given. With the new structure for a split-plot design, some adaptation and expansion of the criteria are needed to consider the presence of two types of experimental units, namely the whole-plot and subplot experimental units. A *good* SPD should seek to balance and simultaneously optimize a number of varied criteria:

1. Result in good fit of the model to the data.
2. Provide good model parameter estimates.
3. Provide a good distribution of predict variance of the response,  $\frac{\text{Var}[\hat{y}(x)]}{\sigma^2}$ , throughout the region of interest.
4. Provide an estimate of “pure” experimental error at both the the whole plot and subplot levels.
5. Give sufficient information to allow for lack of fit test.



**TABLE 10.6** A Restricted Box–Behnken Split-Plot Design with Nine Whole Plots for Two Hard-to-Change and Two Easy-to-Change Factors

| Whole Plot # | $z_1$ | $z_2$ | $x_1$    | $x_2$    | # Runs  |
|--------------|-------|-------|----------|----------|---------|
| 1            | -1    | -1    | 0        | 0        | 1       |
| 2            | -1    | +1    | 0        | 0        | 1       |
| 3            | +1    | -1    | 0        | 0        | 1       |
| 4            | +1    | +1    | 0        | 0        | 1       |
| 5            | -1    | 0     | $\pm 1$  | 0        | 4       |
|              |       |       | 0        | $\pm 1$  |         |
| 6            | +1    | 0     | $\pm 1$  | 0        | 4       |
|              |       |       | 0        | $\pm 1$  |         |
| 7            | 0     | -1    | $\pm 1$  | 0        | 4       |
|              |       |       | 0        | $\pm 1$  |         |
| 8            | 0     | +1    | $\pm 1$  | 0        | 4       |
|              |       |       | 0        | $\pm 1$  |         |
| 9            | 0     | 0     | $\pm 1$  | $\pm 1$  | 4 + #CR |
|              |       |       | <b>0</b> | <b>0</b> |         |

6. Provide a check on the homogeneity of variance in both the whole plot and subplot portions of the design space.
7. Be insensitive (robust) to the presence of outliers in the data.
8. Be robust to errors in the control of design levels.
9. Consider the relative cost in setting hard-to-change and easy-to-change factors.
10. Consider simplicity and symmetry of the design.
11. Consider the required complexity of the analysis.
12. Be cost-effective.

Properties 4, 6, 9, and 11 have been adapted from the list in Section 8.1 to consider the special features of split-plot designs. As with completely randomized designs, not all of these properties are required in every split-plot application and no single design is capable of performing best in all of these properties simultaneously. However, it is beneficial when choosing a designed experiment to give serious consideration to the relative importance of these criteria. Often a good design can do quite well for several of these properties.

Item 4 has been expanded to note that the two error terms now need to both be estimated. Frequently obtaining a good estimate of the subplot error term is relatively easy, but because of the nature of hard-to-change factors, many designs do not place adequate emphasis of estimating the whole plot error term. Similarly for Item 6, it may be difficult to obtain sufficient information about homogeneity of variance in the whole plot space. Item 9 deserves special emphasis as the relative cost of setting the hard-to-change and easy-to-change factors will typically dictate what designs can be considered. Some additional discussion of this is presented subsequently. Finally, Item 11 has added importance in the split plot setting as some alternatives exist for the estimation and analysis, depending on the type of design selected. Parker et al. (2008) present a thorough discussion about how to balance several criteria in the split-plot design case.

**Example 10.4 The Adhesive Strength Experiment** Consider an experiment involving four factors that potentially affect the strength of an adhesive intended for application in a medical device application. Two of the factors ( $A =$  cure temperature and  $B =$  percent of resin in the adhesive) are relatively hard to change from run to run, while the other two ( $C =$  amount of adhesive applied and  $D =$  cure time) are easy to change. A second-order model is thought to be appropriate for the underlying relationship between strength and the input factors. The form of the model is

$$\begin{aligned} \text{Strength} = & \beta_0 + \beta_1 A + \beta_2 B + \beta_{12} AB + \beta_{11} A^2 + \beta_{22} B^2 + \gamma_1 C + \gamma_2 D \\ & + \gamma_{12} CD + \gamma_{11} C^2 + \gamma_{22} D^2 + \alpha_{11} AC + \alpha_{12} AD + \alpha_{21} BC \\ & + \alpha_{22} BD + \delta + \varepsilon \end{aligned}$$

where the  $\beta$  terms involve hard-to-change factors, the  $\gamma$  terms involve easy-to-change factors and the  $\alpha$  terms involve both hard- and easy-to-change factors. Written in matrix form, we have a model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon}$$

with  $\mathbf{X}$  is the  $N \times p$  design matrix expanded to model form for the  $p$  model parameters which include the intercept;  $\mathbf{Z}$  is an  $N \times a$  classification matrix of ones and zeros where the  $ij$ th entry is 1 if the  $i$ th observation ( $i = 1, \dots, N$ ) belongs to the  $j$ th whole plot ( $j = 1, \dots, a$ ). Note the two errors terms:  $\boldsymbol{\delta} \sim N(0, \sigma_{WP}^2)$  for the whole plot and  $\boldsymbol{\varepsilon} \sim N(0, \sigma^2)$  for the subplot error. We also assume  $\boldsymbol{\delta}$  and  $\boldsymbol{\varepsilon}$  are mutually independent.

The covariance matrix of the responses in a split-plot design is

$$\begin{aligned} \text{Var}(\mathbf{y}) &= \boldsymbol{\Sigma} = \sigma_{WP}^2 \mathbf{Z}\mathbf{Z}' + \sigma^2 \mathbf{I}_N \\ &= \sigma^2 [d\mathbf{Z}\mathbf{Z}' + \mathbf{I}_N] \end{aligned}$$

where  $\mathbf{I}_N$  is an  $N \times N$  identity matrix and  $d = \sigma_{WP}^2 / \sigma^2$  represents the variance component ratio. This type of error structure is called compound symmetric. If we sort the observations by whole plots, we can write  $\boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_a)$  where each  $n_j \times n_j$  matrix  $\boldsymbol{\Sigma}_j$  is given by

$$\boldsymbol{\Sigma}_j = \begin{bmatrix} \sigma_{WP}^2 + \sigma^2 & \dots & \sigma_{WP}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{WP}^2 & \dots & \sigma_{WP}^2 + \sigma^2 \end{bmatrix}$$

and  $\boldsymbol{\Sigma}_j$  denotes the covariance matrix of responses for the  $j$ th whole plot. Note that the variance of an individual observation is the sum of the subplot and whole plot error variances,  $\sigma_{WP}^2 + \sigma^2$ .

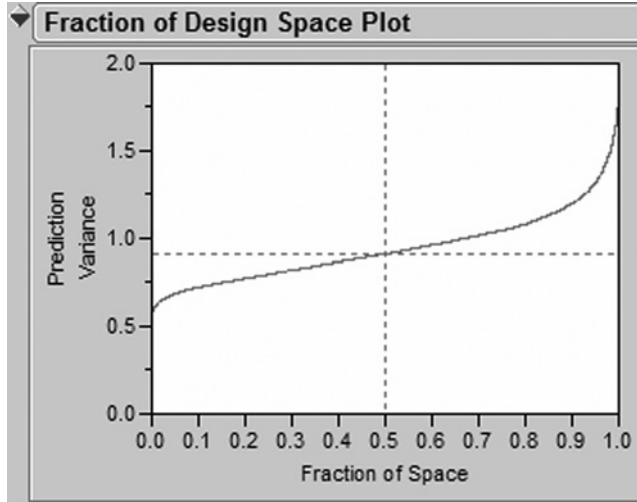
**TABLE 10.7 The Response Surface Design and Data for the Adhesive Strength Example**

| Whole Plot # | A  | B  | C  | D  | Strength |
|--------------|----|----|----|----|----------|
| 1            | 1  | -1 | 1  | -1 | 15.7     |
|              |    |    | -1 | 1  | 27.0     |
| 2            | -1 | 1  | -1 | 1  | 46.8     |
|              |    |    | 1  | -1 | 29.9     |
| 3            | -1 | -1 | -1 | 1  | 17.6     |
|              |    |    | 1  | -1 | 45.9     |
| 4            | -1 | 1  | 1  | 1  | 36.8     |
|              |    |    | -1 | -1 | 32.1     |
| 5            | -1 | -1 | 1  | 1  | 30.7     |
|              |    |    | -1 | -1 | 53.7     |
| 6            | 1  | 1  | 1  | -1 | 8.1      |
|              |    |    | -1 | 1  | 45.2     |
| 7            | 1  | 0  | -1 | -1 | 19.4     |
|              |    |    | 0  | 0  | 33.6     |
| 8            | 1  | -1 | -1 | -1 | 18.3     |
|              |    |    | 1  | 1  | 19.6     |
| 9            | 0  | -1 | 0  | -1 | 42.9     |
|              |    |    | -1 | 0  | 42.3     |
| 10           | 1  | 1  | 1  | 1  | 59.1     |
|              |    |    | -1 | -1 | 16.1     |
| 11           | 0  | 1  | 0  | 0  | 45.3     |
|              |    |    | 1  | -1 | 31.6     |
| 12           | 0  | 0  | 0  | 1  | 46.7     |
|              |    |    | 1  | 0  | 46.2     |

The experimenters have a budget that will allow them to reset the hard-to-change factors 12 times, with 2 runs of different levels of the easy-to-change factors within each whole plot. Table 10.7 shows the *I*-optimal design satisfying the experimenters logistical constraints generated in JMP with the resulting data obtained from the experiment. As with completely randomized designs it is beneficial to explore the prediction variance properties of the constructed design for the assumed model. Figure 10.11 shows the range of prediction variance values throughout the design region in the fraction of design space plot generated in JMP. Liang et al. (2006a) provide additional alternatives to this plot to allow for more detailed examination of the prediction performance throughout the design space. For interactive exploration of the prediction variance, the prediction profiler described in Chapter 9 and shown in Fig. 10.12 allows for better understanding.

Because the observations are not independent and have a compound symmetric error structure generalized least squares (GLS) is required to obtain estimates of the model parameters. The software package JMP uses the REML (REstricted or REsidual Maximum Likelihood) method of fitting mixed models (those which contain both fixed effects—the factor effects, and random effects—the whole-plot error terms). See Wolfinger et al. (1994) and Searle et al. (1992) for more details on this estimation approach.

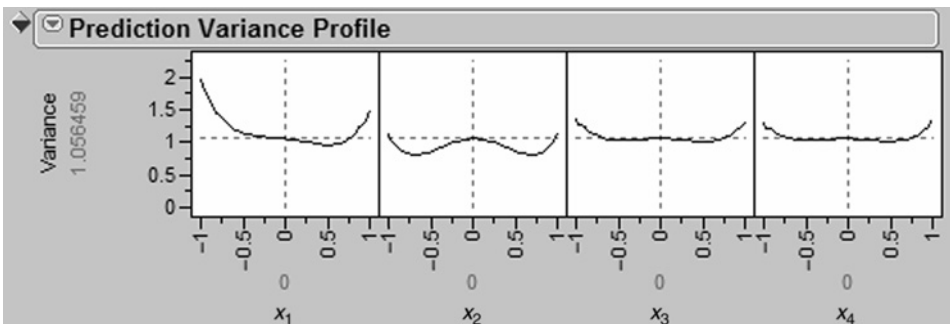
Table 10.8 shows the output generated by JMP for the adhesive strength data. Notice that the whole plot variance component is much larger than the subplot variance component, as is typically the case in split-plot designs. The whole plot variance component accounts for about 70% of the total variability.



**Figure 10.11** The fraction of design space plot for the adhesive strength example.

**Design Considerations for Split-Plot Experiments** Because of the more complex structure for split-plot designs, there are some additional considerations for selecting a design. Adaptations of the central composite and Box–Behnken designs described in Chapter 8 are a popular choice for this setting, since many of their desirable properties continue to exist. Much depends on the experiment constraints as in the split-plot there are many aspects to consider.

**Optimality** As with completely randomized designs, computer generated designs are available to optimize based on various criteria. If interest lies primarily in estimating the model parameters, then the *D*-optimal criterion described in Section 9.2 can be adapted to the split-plot setting. The optimal design is strictly dependent on the ratio of the two variance components,  $d = \sigma^2_{WP} / \sigma^2$ , although in many situations the design performance is quite robust to a wide range of values of this ratio. Typically this ratio ranges from close to zero, if there is little additional variability associated with the whole plots, to values as large as 10 or 20. Goos (2002) developed methods for generating designs for a variety of specified



**Figure 10.12** The prediction profiler for the adhesive strength example.

**TABLE 10.8 Output from JMP for Adhesive Strength Example**

| Response Strength Summary of Fit |          |
|----------------------------------|----------|
| R-Square                         | 0.969691 |
| R-Square Adj                     | 0.922543 |
| Root Mean Square Error           | 4.369115 |
| Mean of Response                 | 44.1125  |
| Observations (or Sum Wgts)       | 24       |

| <i>Parameter Estimates</i> |           |           |       |         |           |  |
|----------------------------|-----------|-----------|-------|---------|-----------|--|
| Term                       | Estimate  | Std Error | DFDen | t-Ratio | Prob >  t |  |
| Intercept                  | 46.630773 | 6.134175  | 5.275 | 7.60    | 0.0005    |  |
| A                          | -5.202121 | 2.582063  | 4.743 | -2.01   | 0.1031    |  |
| B                          | 3.2464223 | 2.366201  | 4.752 | 1.37    | 0.2313    |  |
| C                          | 0.7165934 | 1.051328  | 3.605 | 0.68    | 0.5367    |  |
| D                          | 3.4042293 | 1.0762    | 3.635 | 3.16    | 0.0389    |  |
| A*A                        | 2.7031269 | 5.595789  | 6.018 | 0.48    | 0.6461    |  |
| A*B                        | 3.125     | 2.637179  | 4.708 | 1.18    | 0.2924    |  |
| B*B                        | -1.594711 | 6.387644  | 5.214 | -0.25   | 0.8124    |  |
| A*C                        | -0.180581 | 1.089195  | 3.464 | -0.17   | 0.8775    |  |
| B*C                        | 0.6198131 | 1.081642  | 3.473 | 0.57    | 0.6017    |  |
| C*C                        | -0.881511 | 3.294511  | 3.627 | -0.27   | 0.8035    |  |
| A*D                        | 1.0819185 | 1.089195  | 3.464 | 0.99    | 0.3848    |  |
| B*D                        | 8.4263892 | 1.085298  | 3.504 | 7.76    | 0.0025    |  |
| C*D                        | 3.4088539 | 2.70862   | 7.694 | 1.26    | 0.2450    |  |
| D*D                        | -2.134799 | 3.192447  | 3.597 | -0.67   | 0.5441    |  |

| <i>REML Variance Component Estimates</i> |           |               |           |           |           |              |
|--|-----------|---------------|-----------|-----------|-----------|--------------|
| Random Effect                            | Var Ratio | Var Component | Std Error | 95% Lower | 95% Upper | Pct of Total |
| Whole Plots                              | 2.4146228 | 46.093136     | 37.686598 | -27.7726  | 119.95887 | 70.714       |
| Residual                                 |           | 19.089166     | 14.600518 | 6.451241  | 205.70398 | 29.286       |
| Total                                    |           | 65.182302     |           |           |           | 100.000      |
| -2 LogLikelihood = 96.742353034          |           |               |           |           |           |              |

constraints. The total number of runs, the number of whole plots, or a combination of both can be specified for a particular experiment. The statistical software, JMP, can generate  $D$ -optimal designs for specified numbers of run and whole plots.

If the primary goal of the experiment is to obtain good prediction throughout the design space, then  $G$ - and  $I$ -optimality can be considered.  $G$ -optimality minimizes the worst prediction in the design space, while  $I$ -optimality minimizes the average prediction. While both criteria are possible, recent work has focused more on  $I$ -optimality since it focuses on the central tendency of prediction variance, rather than just its extreme. As with  $D$ -optimal designs, the choice of an  $I$ -optimal design is related to the choice of  $\sigma_{WP}^2/\sigma^2$ , but generally quite robust over a range of values. The statistical software, JMP, can also generate  $I$ -optimal designs for specified numbers of run and whole plots.

Using either  $D$ -optimality or  $I$ -optimality, it is possible to generate flexible designs that meet the needs of particular experimental design situations. Alternately, it is possible to adapt the levels of a central composite design to optimize it for a split-plot design. For the restricted design of Table 10.5, different values of the factorial portion of the design are appropriate for the hard-to-change and easy-to-change factors to optimize design performance based on  $D$ -,  $G$ -, or  $I$ -optimality. The choice of which values should be selected is dependent on the criteria considered as well as the ratio of the two variance components. For more details on this approach, see Liang et al. (2006b).

**Balance** If changing the hard-to-change factors settings is cost- or time-intensive relative to the easy-to-change factors settings, it may be sensible to run as many subplot runs per whole plot as possible. This will allow for maximum use to be made of each reset of the hard-to-change factors. For example, in the polyamide resin example (Example 8.5), consider the scenario where the temperature at the time of addition required considerable time to reset since this temperature was required to stabilize before experimental units could be run. The other two factors, agitation and rate of addition, could be quickly and easily reset. In this case, it may be sensible to consider a balanced split-plot design. A **balanced** split-plot design is one that has equal number of subplot runs for each whole plot.

The restricted central composite design of Table 10.5 and the restricted Box–Behnken design of Table 10.6 are extreme examples of unbalanced split-plot designs, since several of the whole plots have just a single run. Vining et al. (2005) and Parker et al. (2007) give several examples of balanced adaptations of central composite and Box–Behnken designs for different numbers of hard-to-change and easy-to-change factors.

**Cost** The total cost of a split-plot experiment is a function of whole plot costs as well as subplot costs. Bisgaard (2000) and Liang et al. (2007) formulate the cost of a split-plot experiment as

$$C = C_{WP}a + C_{SP}N$$

where  $C$  denotes the total cost of the experiment,  $a$  denotes the number of whole plot units,  $N$  is the total number of observations.  $C_{WP}$  and  $C_{SP}$  are the costs associated with the individual whole plots and subplots, respectively. Note that the cost incurred for measuring the response for each observation is considered a part of subplot costs.

In practice, it may be difficult to estimate the exact costs associated with whole plots or subplots (i.e., precise values for  $C_{WP}$  and  $C_{SP}$ ). It may be more feasible to specify the

relative cost of these quantities (i.e.,  $r = C_{SP}/C_{WP}$ ). Hence, the cost of the experiment is proportional to  $a + rN$  with

$$C \propto a + rN.$$

Typically we might expect  $C_{WP}$  to be greater than  $C_{SP}$  since this involves the hard-to-change factors where more time/effort is required to change levels of the whole plot factors. As  $C_{WP}$  increases relative to  $C_{SP}$ ,  $r$  approaches zero. On the other hand, if obtaining the measurement of the response for each observation is expensive, then  $C_{SP}$  may become larger compared to  $C_{WP}$  and  $r$  will increase.

The  $D$ -,  $G$ -, and  $I$ -optimality criteria from Chapter 8 can be adapted for the split-plot setting using this cost structure. For  $D$ -optimality, the moment matrix for split-plot designs is  $(\mathbf{X}'\Sigma^{-1}\mathbf{X})$ . We can define the cost-adjusted moment matrix for split-plot designs as

$$M = \frac{(\sigma_{WP}^2 + \sigma^2)(\mathbf{X}'\Sigma^{-1}\mathbf{X})}{C}$$

For  $G$ - and  $I$ -optimality, cost-adjusted prediction variance (CPV) is defined as

$$\text{CPV} = \frac{C \text{Var}(\hat{y}(x_0))}{\sigma_{WP}^2 + \sigma^2}$$

It is noteworthy that we can think of the completely randomized design as a special case of a split-plot design where each observation can be treated as a separate whole plot (in this case,  $a = N$ ). In CRDs the number of whole plots and subplots are equal and the total cost of the experiment is given by

$$C = (C_{WP} + C_{SP})N = C_{WP}(1 + r)N$$

This expression is proportional then to the standard penalty of  $N$  commonly used for the  $D$ -,  $G$ -, and  $I$ -optimality criteria in completely randomized design. Parker et al. (2008) give a detailed comparison of several designs and the trade-offs between different criteria for a variety of cost and variance component ratios. Lu and Anderson-Cook (2014) and Lu, Robinson, and Anderson-Cook (2014) discuss multiple criterion optimization for split-plot designs with cost treated as one or two of these criteria.

**Equivalent Estimation** Split-plot designs can also be constructed such that ordinary least squares (OLS) is equivalent to generalized least squares (GLS) for estimation. Initially this may seem like an unintuitive goal for constructing a design, since the correct analysis approach based on the model is GLS. However, the benefits of this class of design are substantial.

- Design performance is independent of the variance components. As a result, no estimate of the ratio of  $\sigma_{WP}^2/\sigma^2$  is required before selecting a design.

**TABLE 10.9** A Balanced Equivalent Estimation Central Composite Design Using the VKM Construction Method for Two Hard-to-Change and Two Easy-to-Change Factors with Ten Whole Plots and Four Runs per Whole Plot

| Whole Plot # | $z_1$    | $z_2$    | $x_1$       | $x_2$       | # Runs |
|--------------|----------|----------|-------------|-------------|--------|
| 1            | -1       | -1       | $\pm 1$     | $\pm 1$     | 4      |
| 2            | -1       | +1       | $\pm 1$     | $\pm 1$     | 4      |
| 3            | +1       | -1       | $\pm 1$     | $\pm 1$     | 4      |
| 4            | +1       | +1       | $\pm 1$     | $\pm 1$     | 4      |
| 5            | $-\beta$ | 0        | 0           | 0           | 4      |
| 6            | $+\beta$ | 0        | 0           | 0           | 4      |
| 7            | 0        | $-\beta$ | 0           | 0           | 4      |
| 8            | 0        | $+\beta$ | 0           | 0           | 4      |
| 9            | 0        | 0        | $\pm\alpha$ | 0           | 4      |
|              |          |          | 0           | $\pm\alpha$ |        |
| 10           | 0        | 0        | 0           | 0           | 4      |

- Model parameter estimates are independent of the variance component estimates. Using REML, the model parameter estimates are typically dependent on the variance components, which may not be well estimated if few degrees of freedom are available for their estimation.
- The designs can be constructed so the equivalence property holds for any model nested in the complete second-order model. When the data are collected and the model parameters estimated, some terms of the model may not be significant and we wish to reduce the model. In this case, the estimates of the new reduced model can still be estimated using OLS.
- The model parameter estimates are the best linear unbiased estimate (BLUE), and this property is robust to the assumption of normality of the error terms.

For more discussion on the advantages of equivalent estimation designs see Parker et al. (2007).

Several construction techniques exist for creating these designs based on the type of response surface design that is used for the starting point, restrictions on numbers of whole plots, size of whole plots and the ability to select different axial values. Table 10.9 shows an adapted balanced Central Composite design for two hard-to-change and two easy-to-change factors with twelve whole plots and four runs per whole plot. It uses the Vining et al. (2005) (VKM) approach which allows for flexible axial values for both the whole and subplot factors. Table 10.10 shows an adapted Box–Behnken design with the same number and type of factors with the minimum number of allowable whole plots. This design uses the minimum whole plot (MWP) method which allows for subplot center runs to be added to achieve the required conditions for equivalence. Parker et al. (2006, 2007) provide examples and details on the construction for a large number of cases.

The replication of subplot combinations within whole plots reflects the assumption that the cost of setting the hard-to-change factors dominates the cost of setting the easy-to-change factors. It has the added benefit that there are many degrees of freedom available for a pure error estimate of the subplot error,  $\sigma^2$ .



**TABLE 10.10** A Balanced Equivalent Estimation Box–Behnken Design Using the MWP Construction Method for Two Hard-to-Change and Two Easy-to-Change Factors with Nine Whole Plots and Five Runs per Whole Plot

| Whole Plot # | $z_1$ | $z_2$ | $x_1$   | $x_2$   | # Runs |
|--------------|-------|-------|---------|---------|--------|
| 1            | -1    | -1    | 0       | 0       | 5      |
| 2            | -1    | +1    | 0       | 0       | 5      |
| 3            | +1    | -1    | 0       | 0       | 5      |
| 4            | +1    | +1    | 0       | 0       | 5      |
| 5            | -1    | 0     | $\pm 1$ | 0       | 5      |
|              |       |       | 0       | $\pm 1$ |        |
|              |       |       | 0       | 0       |        |
| 6            | +1    | 0     | $\pm 1$ | 0       | 5      |
|              |       |       | 0       | $\pm 1$ |        |
|              |       |       | 0       | 0       |        |
| 7            | 0     | -1    | $\pm 1$ | 0       | 5      |
|              |       |       | 0       | $\pm 1$ |        |
|              |       |       | 0       | 0       |        |
| 8            | 0     | +1    | $\pm 1$ | 0       | 5      |
|              |       |       | 0       | $\pm 1$ |        |
|              |       |       | 0       | 0       |        |
| 9            | 0     | 0     | $\pm 1$ | $\pm 1$ | 5      |
|              |       |       | 0       | 0       |        |
|              |       |       | 0       | 0       |        |

**10.8 RSM FOR NON-NORMAL RESPONSES—GENERALIZED LINEAR MODELS**

Generalized Linear Models are a unified collection of models that accommodate response distributions that obey the **exponential family**. These response distributions include the normal, Poisson, binomial, exponential, and gamma. The collection of models encompassed by the GLMs include both linear and nonlinear models. The common thread is the use of a **linear predictor**  $\mathbf{x}'\beta$ , where the vector  $\mathbf{x}$  is in model form. For example, important members of the class of GLMs are given by the **exponential model**

$$E(y) = e^{\mathbf{x}'\beta} \tag{10.49}$$

and the **inverse polynomial model**

$$E(y) = \frac{1}{\mathbf{x}'\beta} \tag{10.50}$$

While both models are nonlinear in the parameters, they contain a linear component and a linear predictor, hence the word “linear” in their name. As one would expect, GLM applications involve using the response surface model (first-order interactive model, second-order model, etc.) as the linear predictor. As we shall illustrate, certain models are appropriate for certain distributions. For example, a special member of the **GLM** family is the standard linear model

$$E(y) = \mathbf{x}'\beta \tag{10.51}$$

which is particularly appropriate for the case where the response is normal.

Many of the theoretical aspects of GLMs are not presented in this text. For further details the reader is referred to McCullagh and Nelder (1989), Myers and Montgomery (1997), and Myers, Montgomery, Vining, and Robinson (2011).

### 10.8.1 Model Framework: The Link Function

The characterization of the model is in terms of what is called a **link function**. The link function describes the transformation on the response mean  $E(y)$  or  $\mu_y$ , which is equated to the linear predictor. In other words, it is the function that *links the mean to the linear predictor*. As examples, consider the linear model, exponential model, and inverse polynomial model mentioned in the introduction to this section.

The exponential model is derived from the log link, that is,

$$\ln \mu_y = \mathbf{x}'\boldsymbol{\beta} \quad (10.52)$$

and thus the model for the mean is

$$\mu_y = e^{\mathbf{x}'\boldsymbol{\beta}} \quad (10.53)$$

The inverse polynomial model comes from the use of the reciprocal link,

$$\mu_y^{-1} = \mathbf{x}'\boldsymbol{\beta} \quad (10.54)$$

and thus the model for the mean is given by

$$\mu_y = \frac{1}{\mathbf{x}'\boldsymbol{\beta}} \quad (10.55)$$

The linear model is derived from the identity link,  $\mu_y = \mathbf{x}'\boldsymbol{\beta}$ , which of course leads to the same expression for the model itself.

We can say then that if we use a link function

$$g(\mu_y) = \mathbf{x}'\boldsymbol{\beta} \quad (10.56)$$

then the model for the mean  $\mu_y$  is given by

$$\mu_y = g^{-1}(\mathbf{x}'\boldsymbol{\beta}) \quad (10.57)$$

One very important link that will be used subsequently is the **logit link** defined by  $g(\mu_y) = \ln[\mu_y/(1 - \mu_y)]$ , resulting in the model

$$\mu_y = g^{-1}(\mathbf{x}'\boldsymbol{\beta}) = \frac{1}{1 + e^{-\mathbf{x}'\boldsymbol{\beta}}}.$$

### 10.8.2 The Canonical Link Function

For each of the members of the exponential family there is a natural **canonical link function**. This does not imply that one is confined to that canonical link. But the reader should understand that this natural link does enjoy certain theoretical as well as practical advantages. At this point we shall give the canonical links for the binomial, Poisson, exponential, gamma, and normal distributions. For the Poisson case the canonical link function is the log link, giving the exponential model. This model is commonly called the **Poisson regression model**. For the binomial case, the canonical link is the **logit link** with the resulting model given by

$$\mu_y = \frac{1}{1 + e^{-\mathbf{x}'\boldsymbol{\beta}}} \quad (10.58)$$

often referred to as the **logistic regression model**. In the case of the logistic regression model,  $\mu_y$  is often the probability  $P$  of an event, say a defective item. Note that the right-hand side of Equation 10.58 is bounded between 0 and 1. If the experiment involves say  $n$  runs at each design point, and the response is the number out of the  $n$  that are defective, then the mean is  $nP$  and the logistic model becomes

$$\mu_y = \frac{n}{1 + e^{-\mathbf{x}'\boldsymbol{\beta}}}.$$

In the Poisson case note that the exponential model  $\mu_y = e^{\mathbf{x}'\boldsymbol{\beta}}$  does not allow negative values for prediction of Poisson counts.

In the use of both the exponential and gamma distributions the canonical link is the **reciprocal link**  $\mu_y^{-1} = \mathbf{x}'\boldsymbol{\beta}$ , resulting in the inverse polynomial model  $\mu_y = 1/(\mathbf{x}'\boldsymbol{\beta})$ .

### 10.8.3 Estimation of Model Coefficients

When we consider the distributions in the exponential family, it becomes apparent that, apart from the normal distribution, the variance is a function of the mean. Since the mean is subject to change through changes in levels of the design variables or regressor levels, the variance of the response will be nonconstant. Thus we have the nonideal conditions described earlier in this chapter. While the observations on the responses are still assumed to be independent, the nonconstant variance and nonlinearity necessitate consideration of an alternative estimation procedure.

The standard operating procedure for estimation of model coefficients in the GLM scenario is the **method of maximum likelihood**. We will not deal with the details of the likelihood for the exponential family; the reader again is referred to McCullagh and Nelder (1989) and Myers, Montgomery, Vining, and Robinson (2011). The methodology used is easily motivated through a discussion of weighted least squares. Consider, for example, the model

$$y_i = \mu(\mathbf{x}_i, \boldsymbol{\beta}) + \varepsilon_i \quad (i = 1, 2, \dots, n) \quad (10.59)$$

or simply

$$y_i = \mu_i + \varepsilon_i \quad (i = 1, 2, \dots, n) \quad (10.60)$$

Here  $\mu_i$  may be either linear or nonlinear in the parameters. Now if  $\text{Var}(y_i)$  is nonconstant and is denoted by

$$\text{Var}(y_i) = v_i \quad (i = 1, 2, \dots, n)$$

and in fact the  $v_i$  are *known*, then the method of weighted least squares involves finding the coefficient vector  $\beta$  that satisfies

$$\text{Min} \sum_{i=1}^n \left[ \frac{y_i - \mu(\mathbf{x}_i, \beta)}{\sqrt{v_i}} \right]^2 \quad (10.61)$$

Here, the  $v_i$  are assumed known, and thus independent of the model coefficients themselves. This is a reasonably straightforward nonlinear weighted least squares procedure and is discussed in detail in Carroll and Ruppert (1988) and Myers, Montgomery, Vining, and Robinson (2011). However, in GLMs the variance  $v_i$  is a function of the mean and hence a function of model parameters. Thus, the one-step use of weighted least squares cannot be accomplished. Rather the procedure is iterative in nature. The resulting procedure, termed **iterative reweighted least squares** (IRWLS), is described below.

A good illustration of the method of iterative reweighted least squares is provided by the case of the Poisson distribution with the log link. The model is given by

$$\mu_i = e^{\mathbf{x}_i' \beta} \quad (10.62)$$

and the variance  $v_i$  is also  $e^{\mathbf{x}_i' \beta}$ . As a result, a weighted least squares with known  $\beta$  in  $v_i$  implies that the differentiation involves the numerator but the denominator remains fixed. Thus the weighted least squares procedure is to solve (after differentiating Equation 9.61 with respect to  $\beta$ )

$$\sum_i (y_i - e^{\mathbf{x}_i' \beta}) \mathbf{x}_i = \mathbf{0}$$

for  $\beta$ . The left side of the above equation is referred to as the **score function**. Solution of this equation for  $\beta$  is certainly not difficult. However, keep in mind that the procedure described can only be viewed from a conceptual point of view, because  $\beta$  is not known. Thus the IRWLS is required to accomplish what is described above in an iterative fashion.

Suppose we denote the variance of the  $i$ th observation as  $v_i(\beta)$ . The procedure is as follows:

- (i) Begin with a starting value  $\mathbf{b}_0$ .
- (ii) Substitute  $\mathbf{b}_0$  into  $v_i(\beta)$ , forming  $v_{i,0}$ .
- (iii) Do weighted least squares with fixed weights  $v_{i,0}$ .
- (iv) Use the estimate  $\mathbf{b}_1$  from (iii), and recalculate the weights, thus forming  $v_{i,1}$ .
- (v) Go back to (iii), using  $v_{i,1}$ .
- (vi) Continue until convergence.

The above method converges to the maximum likelihood estimator for the coefficient vector  $\beta$ . See McCullagh and Nelder (1989) and Myers, Montgomery, Vining, and Robinson (2011) for details.

#### 10.8.4 Properties of Model Coefficients

As in any maximum likelihood procedure, the model coefficients are asymptotically unbiased. The variance–covariance matrix, which is used for tests on coefficients and thus is of primary importance in variable screening, is not generally free of parameters in  $\beta$ . For example, in the case of logistic and Poisson regression with canonical links, asymptotically

$$\text{Var}(\mathbf{b}) = (\mathbf{X}'\mathbf{V}\mathbf{X})^{-1} \quad (10.63)$$

where  $\mathbf{V}$  is the diagonal variance–covariance matrix of the response observations,

$$\mathbf{V} = \text{diag}\{v_i(\beta)\}$$

Thus for logistic regression and Poisson regression with the log link, we have respectively,  $v_i(\beta) = n_i P_i(1 - P_i)$  with  $P_i = 1/(1 + e^{-x_i'\beta})$  and  $v_i(\beta) = e^{x_i'\beta}$ . Thus any statistical inference that involves standard errors of coefficients makes use of estimated standard errors with  $\mathbf{b}$  substituted for  $\beta$ . Variable screening through tests on single coefficients makes use of so-called **Wald inference**. The ratio

$$\left[ \frac{b_j}{se(b_j)} \right]^2$$

is asymptotically  $\chi^2_1$ , where  $se(b_j)$  is the standard error of  $j$ th model parameter. Thus the  $\chi^2$  test statistic replaces the  $t$ -statistic used in standard least squares methodology. These Wald  $\chi^2$  statistics will be illustrated in examples.

#### 10.8.5 Model Deviance

In certain facets of RSM, tests for lack of fit are important, and residual plots can also be beneficial in evaluating a fitted model. In the case of classical RSM, employed with linear models with constant variance, standard  $F$ -tests are used for lack of fit. In the GLM these  $F$ -tests are not applicable. In classical RSM either ordinary residuals or studentized residuals are plotted. These residuals do not enjoy the same usefulness in a GLM situation, because their values depend critically on the assumption of constant variance. In fact, the broad notion of sums of squares (including regression and residual sums of squares) plays no role of substance in the GLM except in the special case of a normal response with an identity link. For the most part they are replaced by quantities that relate to the concept of **model deviance**. Model deviance is an important component of **likelihood inference**. Likelihood inference is rooted in the fact that, in general, a model is evaluated on the

basis of the value of the log likelihood, **computed as a function of model estimates**. For example, likelihood ratio inference allows the use of the likelihood ratio

$$-2 \ln \left[ \frac{L(\mathbf{b}_r)}{L(\mathbf{b}_f)} \right]$$

where  $L(\mathbf{b}_r)$  is the likelihood computed for a restricted model whereas  $L(\mathbf{b}_f)$  is the likelihood computed for the full model. The above likelihood ratio statistic is asymptotically  $\chi^2_{\Delta}$ , where  $\Delta$  is the difference in the number of model parameters for the two models. Note that this test statistic essentially determines whether the full model results in significantly higher log likelihood. In the case of a linear model with normal responses, the log likelihood is, in fact, apart from constants, the residual sum of squares.

The model deviance is technically a lack-of-fit test in which the model in question is compared with the **saturated model**. The saturated model is that which has zero degrees of freedom for residual. That is, we write the saturated model as  $y_i = \mu_i + \varepsilon_i$ , where  $\mu_i$  contains no regressors. Rather there are  $N$  distinct means and thus there are  $N$  parameters to be estimated. In fact, in the case of a binomial response, the  $i$ th binomial observation  $y_i$  is the estimator of the binomial parameter  $P_i$ . In the case of a nongrouped, or **Bernoulli**, scenario, the estimator of  $P_i$  is a zero or one. Formally, then, the calculated model deviance is given by

$$\lambda(\mathbf{b}) = -2 \ln \left[ \frac{L(\mathbf{b})}{L(\hat{\boldsymbol{\mu}})} \right] \quad (10.64)$$

where  $L(\mathbf{b})$  is the maximum value of the likelihood for the model in question and  $L(\hat{\boldsymbol{\mu}})$  is the maximum value of the likelihood for the saturated model. Asymptotically  $\lambda(\mathbf{b})$  is  $\chi^2_{N-p}$ , where  $p$  is the number of parameters in the fitted model in question. Formally, an insignificant value of  $\lambda(\mathbf{b})$  in an upper-tail one-tailed test implies that the fit of the model is not significantly worse than that of the saturated model. Thus, a relatively small value of deviance is favorable for the fitted model. More discussion will be given later regarding the deviance value as a goodness-of-fit statistic. Often the rule of thumb is applied that the quality of the fit is reasonable if  $\lambda(\mathbf{b})/(N-p)$  is not appreciably larger than 1. This rule is prompted by the fact that  $N-p$  is the mean of the  $\chi^2_{N-p}$  distribution.

There are instances in which the asymptotic result here is not satisfactory for small samples. For details, see Myers, Montgomery, Vining, and Robinson (2011).

Tests of nested hypotheses using the likelihood ratio criterion can be accomplished with *differences in deviance*, much like the use of error sums of squares for the case of linear models. Note that from the definition of deviance in Equation 10.64 a likelihood ratio test statistic  $-2 \ln[L(\text{reduced})/L(\text{full})]$  can be written as  $\lambda(\text{reduced}) - \lambda(\text{full})$ , since  $L(\hat{\boldsymbol{\mu}})$  in Equation 10.64 cancels out in the difference in deviances. For example, in a case with four parameters  $\beta_0, \beta_1, \beta_2, \beta_3$  where there is interest in testing  $H_0: \beta_1 = \beta_2 = 0$ , the likelihood ratio statistic can be written

$$2 \ln L(b_0, b_1, b_2, b_3) - 2 \ln L(b_0^*, b_3^*) = \lambda(b_0^*, b_3^*) - \lambda(b_0, b_1, b_2, b_3).$$

In fact, deviance can be used quite simply in a stepwise algorithm. The use of deviance and thus log likelihood in nested hypotheses is not as fragile in small samples as the use of deviance as a goodness-of-fit test.

### 10.8.6 Overdispersion

The concept of overdispersion is extremely important for complete understanding of the practical aspects of the fitting of generalized linear models. The term **overdispersion** implies that variability exists over and above the natural variability associated with the distribution in question. Even if the distributional assumption is correct, there may be **extra** binomial variability or extra Poisson variability. In other words, we say that the **scale parameter**  $\sigma^2 > 1$  is such that the variance of an individual observation is  $\sigma^2 np(1-p)$  rather than  $np(1-p)$ . In the Poisson case with mean  $\lambda$ , this means that the variance is  $\lambda \cdot \sigma^2$ .

There are reasonable explanations for overdispersion. Details regarding its possible sources appear in Myers, Montgomery, Vining, and Robinson (2011). One may view the overdispersion scale parameter  $\sigma^2$  as playing the same role as the error variance in standard linear models. Certainly, sloppy or inadequate modeling in linear models increases the estimate of the error variance, since bias due to lack of fit inflates the estimate. The same is true in the case of the estimator of  $\sigma^2$  in the GLM. Here the **mean deviance**, that is, deviance/( $N-p$ ), plays the role of the mean square error in linear models. Overdispersion often is a result of nonhomogeneity of experimental units. For example, a Poisson-distributed response may be subject to overdispersion when the experimental materials used are from different sources.

**Effect of Overdispersion on Results** The effect of overdispersion matches very closely the effect of nonhomogeneous experimental units in linear models on the error variance. The result is to increase the standard errors of estimates of model coefficients. For example, in the case of models using the canonical link (e.g., logistic regression or Poisson regression with an exponential model), the variance–covariance matrix of coefficients is given by

$$\text{Var}(\mathbf{b}) = (\mathbf{X}'\mathbf{V}\mathbf{X})^{-1}\sigma^2 \quad (10.65)$$

and thus the standard errors are underestimated, since the parameter  $\sigma^2 > 1$  will be ignored in the computations. Obviously, in biological applications involving the use of animals as experimental units, some overdispersion is expected. Whether or not overdispersion is expected in an industrial application depends on the nature of the process.

**Adjustments for Overdispersion** It is standard procedure to adjust estimated standard errors to allow for overdispersion. As one might expect, an estimate of the scale parameter is given by

$$\hat{\sigma}^2 = \frac{\text{deviance}}{N-p} \quad (10.66)$$

which is analogous to the error mean square in standard linear models. A second estimate is given by the mean square of the **Pearson  $\chi^2$  residuals**. This statistic is given by

$$\frac{1}{N-p} \sum_{i=1}^N \left[ \frac{(y_i - \hat{y}_i)^2}{\hat{v}_i} \right] = \frac{\chi^2}{N-p} \tag{10.67}$$

The intuition here is quite clear. The  $v_i$  reflects the natural variance that comes from the type of distribution based on the analysis. Clearly then the quantity inside brackets above is the square of a standardized residual. The expression  $\chi^2/(N-p)$  estimates unity in a non-overdispersion scenario. It should be clear then that when overdispersion exists, an adjustment to the computed standard error is made by multiplying the ordinary standard errors by the factor  $\sqrt{\chi^2/(N-p)}$ , or  $\sqrt{\text{deviance}/(N-p)}$ . SAS PROC GENMOD, a software package widely used in fitting the GLM, allows for the adjustment and will be illustrated with an example in Section 10.8.7.

**10.8.7 Examples**

In this section we present two examples: The first represents a designed experiment containing a binomial response, and the second involves a Poisson response. For the binomial example a logistic regression model is fitted to the data, while for the Poisson example the canonical or log link is used, implying an exponential model.

**Example 10.5 Survival of Spermatozoa** This example is a spermatozoa survival study. The endpoint, or response variable in the experiment, is survival or nonsurvival. The spermatozoa are stored in sodium citrate and glycerol. The amounts of these substances were varied along with a third factor, equilibration time, in a  $2^3$  factorial design. Fifty samples of the material were subjected to the conditions of each of the eight design points. Interpretation of the effects of the factors was sought. The data appear in Table 10.11.

**TABLE 10.11 Spermatozoa Survival Data**

| Sodium Citrate<br>$x_1$ | Glycerol<br>$x_2$ | Equilibrium Time<br>$x_3$ | Number Surviving<br>$y$ |
|-------------------------|-------------------|---------------------------|-------------------------|
| -1                      | -1                | -1                        | 34                      |
| 1                       | -1                | -1                        | 20                      |
| -1                      | 1                 | -1                        | 8                       |
| 1                       | 1                 | -1                        | 21                      |
| -1                      | -1                | 1                         | 30                      |
| 1                       | -1                | 1                         | 20                      |
| -1                      | 1                 | 1                         | 10                      |
| 1                       | 1                 | 1                         | 25                      |



The full model to be fitted to the data contains three linear main effects and all two-factor interactions. The binomial random variable  $y$  is the response. The logit link results in the logistic regression model

$$E(y) = P = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3)}}$$

where  $P$  is the probability of survival.

At this point we will give the likelihood inference designed to test the adequacy of the entire model. This is analogous to the  $F$ -test used in standard linear models. For the intercept-only model

$$P = \frac{1}{1 + e^{-\beta_0}}$$

we have (from PROC GENMOD)

$$-2 \ln L(\text{reduced}) = 544.234$$

whereas for the full model containing seven parameters,

$$-2 \ln L(\text{full}) = 496.055$$

Thus the appropriate likelihood ratio test statistic is

$$\begin{aligned} \chi^2 &= 2 \ln \left[ \frac{L(\text{reduced})}{L(\text{full})} \right] \\ &= 544.234 - 496.055 \\ &= 48.179 \end{aligned}$$

with  $7 - 1 = 6$  degrees of freedom. The value is significant at  $P < 0.0001$ . This implies that the use of the full model significantly *increases* the log likelihood over that of the intercept-only model. The maximum likelihood estimates, standard errors, and Wald  $\chi^2$  values for each coefficient appear in Table 10.12.

**TABLE 10.12** Maximum Likelihood Estimates and Wald Inference on Individual Coefficients for Data of Example 10.5

| Model Term | df | Parameter Estimate | Standard Error | Wald $\chi^2$ | $P$ -Value |
|------------|----|--------------------|----------------|---------------|------------|
| Intercept  | 1  | -0.3770            | 0.1101         | 11.7198       | 0.0006     |
| $x_1$      | 1  | 0.0933             | 0.1101         | 0.7175        | 0.3970     |
| $x_2$      | 1  | -0.4632            | 0.1101         | 17.7104       | 0.0001     |
| $x_3$      | 1  | 0.0259             | 0.1092         | 0.0563        | 0.8124     |
| $x_1 x_2$  | 1  | 0.5851             | 0.1101         | 28.2603       | 0.0001     |
| $x_1 x_3$  | 1  | 0.0544             | 0.1093         | 0.2474        | 0.6189     |
| $x_2 x_3$  | 1  | 0.1122             | 0.1088         | 1.0624        | 0.3027     |

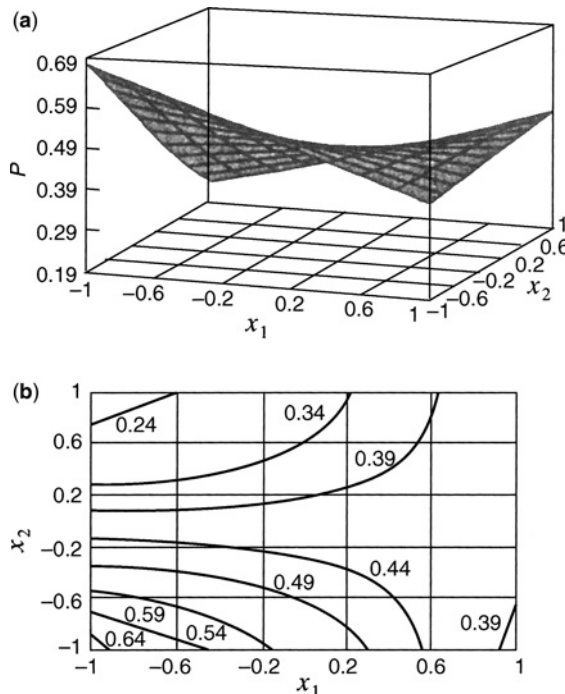
**TABLE 10.13    The Reduced Model for Example 10.5**

| Model Term | df | Parameter Estimate | Standard Error | $\chi^2$ | P-Value |
|------------|----|--------------------|----------------|----------|---------|
| Intercept  | 1  | -0.3637            | 0.1081         | 11.3116  | 0.0008  |
| $x_2$      | 1  | -0.4505            | 0.1084         | 17.2677  | 0.0001  |
| $x_1x_2$   | 1  | 0.5747             | 0.1086         | 27.9923  | 0.0001  |

Note that  $\beta_0$ ,  $\beta_2$ , and  $\beta_{12}$  appear to be important model terms. Any type of stepwise algorithm using log likelihood results in the choice of the same model terms. The reduced model was fitted to the data, and the results are shown in Table 10.13. Thus the fitted model for estimating probability of survival is given by

$$\hat{P} = \frac{1}{1 + e^{-(-0.3637 - 0.4505x_2 + 0.5747x_1x_2)}}$$

Note that fitting the reduced model results in different coefficients than those that appeared in the full model. This is traced to the fact that, unlike the case of *linear* response surface models, for this model the  $2^3$  factorial does not enjoy the property of orthogonality. Also note that the standard errors are not equal as they would be in the case of a linear model with standard assumptions. The contour plot and response surface plot of this model are shown in Fig. 10.13.



**Figure 10.13** Response surface and contour plot for Example 10.5. (a) Response surface. (b) Contour plot.

From the above fitted model and the plots in Fig. 10.13 it is easily established that when the sodium citrate is at the low level, the amount of glycerol has a negative effect on the probability of survival. By that we mean that a low level of glycerol produces the desired result ( $\hat{P}$  close to 1.0). On the other hand, if the sodium citrate level is high, the amount of glycerol has very little effect. The most desirable result occurs when both  $x_1$  and  $x_2$  are simultaneously at the low level, whereas the worst case in terms of survival occurs when sodium citrate is low and glycerol is high.

One must keep in mind that, as in any designed experiment, the conclusions drawn here depend greatly on the chosen ranges of the design variables. In addition, standard techniques such as residual analysis and the use of other diagnostic information should be considered. We will discuss this later.

We now consider **odds ratios** and what they mean in the context of Example 10.5. For the case where a logistic regression model is fitted, the type of interpretation that one obtains from the factor effects that are used for linear models are replaced by odds ratios. In the case of linear models, effect calculations discussed in Chapters 3 and 4 as the difference between the mean response at the high level and that at the low level of a factor are quite natural, simply because the model is in fact linear. In the case of the logistic model the type of interpretation that is natural exploits the *multiplicative* component of the model. In this example we define the **odds** of survival as  $P/(1-P)$ , where, of course,  $P$  is the probability of survival.

From the logistic model in Equation 10.58 we can write

$$\frac{P}{1-P} = e^{x'\beta}$$

The multiplicative nature of the above equation leads to a natural interpretation of the effect of a factor as the *ratio* (rather than the difference) of the odds when a factor is at  $x = +1$  to the odds when the factor is, say, at the midpoint of the range, i.e.,  $x = 0$ —that is, the computation of the ratio of change in the odds of survival as the factor increases one unit. If there is no interaction in the model,  $e^{\beta_2}$  becomes the ratio of the odds of survival at  $x_2 = 1$  to the odds of survival at  $x_2 = 0$ . So the interpretation is quite straightforward in the no-interaction model. In fact,  $(e^{\beta_2})^2 = e^{2\beta_2}$  is the ratio of the odds of survival at  $x_2 = 1$  to the odds at  $x_2 = -1$ . However, when there is interaction in the model, as in Example 10.5,  $e^{\beta_2}$  is the ratio of the odds of survival at  $x_2 = 1$  to the odds at  $x_2 = 0$  when  $x_1 = 0$ . Here  $e^{\beta_2} = 0.637$ , which makes a strong case for not storing the sperm with glycerol at its high value, particularly when  $x_1 = 0$ . For  $x_1 = -1$  the relevant calculated odds ratio for  $x_2$  is  $e^{b_2 - b_{12}} = 0.36$ . This number illustrates how operating at the high level is even more destructive as far as survival is concerned if sodium citrate, is at the low level. On the other hand, if sodium citrate is at the high level, that is,  $x_1 = +1$ , the calculated odds ratio effect of glycerol is  $e^{b_2 + b_{12}} = e^{0.125} = 1.136$ . This odds ratio near one implies that the effect of  $x_2$  (glycerol) is rendered very small when sodium citrate is used at the high level. So, while the maximum estimated probability of survival is achieved with both factors simultaneously at the low level, the use of sodium citrate at the high level negates the effect of glycerol.

It turns out that for all GLMs there is a connection to the effects used in standard linear models. This will be evident in the next example.

**TABLE 10.14**    **Wave Solder Data**

| Run | Factor |    |    |    |    |    |    | y   |     |     |
|-----|--------|----|----|----|----|----|----|-----|-----|-----|
|     | A      | B  | C  | D  | E  | F  | G  | 1   | 2   | 3   |
| 1   | -1     | -1 | -1 | -1 | -1 | -1 | -1 | 13  | 30  | 26  |
| 2   | -1     | -1 | -1 | +1 | +1 | +1 | +1 | 4   | 16  | 11  |
| 3   | -1     | -1 | +1 | -1 | -1 | +1 | +1 | 20  | 15  | 20  |
| 4   | -1     | -1 | +1 | +1 | +1 | -1 | -1 | 42  | 43  | 64  |
| 5   | -1     | +1 | -1 | -1 | +1 | -1 | +1 | 14  | 15  | 17  |
| 6   | -1     | +1 | -1 | +1 | -1 | +1 | -1 | 10  | 17  | 16  |
| 7   | -1     | +1 | +1 | -1 | +1 | +1 | -1 | 36  | 29  | 53  |
| 8   | -1     | +1 | +1 | +1 | -1 | -1 | +1 | 5   | 9   | 16  |
| 9   | +1     | -1 | -1 | -1 | +1 | +1 | -1 | 29  | 0   | 14  |
| 10  | +1     | -1 | -1 | +1 | -1 | -1 | +1 | 10  | 26  | 9   |
| 11  | +1     | -1 | +1 | -1 | +1 | +1 | +1 | 28  |     | 19  |
| 12  | +1     | -1 | +1 | +1 | -1 | +1 | -1 | 100 | 129 | 151 |
| 13  | +1     | +1 | -1 | -1 | -1 | +1 | +1 | 11  | 15  | 11  |
| 14  | +1     | +1 | -1 | +1 | +1 | -1 | -1 | 17  | 2   | 17  |
| 15  | +1     | +1 | +1 | -1 | -1 | -1 | -1 | 53  | 70  | 89  |
| 16  | +1     | +1 | +1 | +1 | +1 | +1 | +1 | 23  | 22  | 7   |

**Example 10.6 Solder Defects**    Consider an electronic circuit card assembly by a wave-soldering process. The response is the number of defects in the solder joint. The process involves baking and preheating the circuit card and passing it through a solder wave by conveyor. Condra (1993) originally presented the results shown in Table 10.14, and Hamada and Nelder (1997) reanalyzed them. The seven factors are (A) prebake condition, (B) flux density, (C) conveyor speed, (D) preheat condition, (E) cooling time, (F) ultrasonic solder agitator, and (G) solder temperature. Each factor is at two levels, and the experimental design is a replicated  $2^{7-3}$  fractional factorial. Note that data point 11 has only two observations. A third observation was reported, but strong evidence suggested that it was an outlier. See Hamada and Nelder (1997).

Table 10.15 displays a SAS GENMOD output for the wave solder data. The response is assumed to obey a Poisson distribution. The log link is used, and thus the model is the exponential model

$$\mu = e^{\mathbf{x}'\boldsymbol{\beta}}$$

with the linear predictor  $\mathbf{x}'\boldsymbol{\beta}$  containing seven main effects and six interaction terms. Note from this output that the mean deviance, deviance/( $N - p$ ), exceeds unity considerably. This signals the condition of overdispersion. The danger is the underestimation of standard errors of coefficients. Also, we see that the mean squared error of the Pearson  $\chi^2$  residuals exceeds unity considerably. Note from the “Analysis of Parameter Estimates” in the output that all main effects are significant apart from  $D$  and  $F$ . The interactions  $AC$ ,  $AD$ ,  $BC$ , and  $BD$  are also statistically significant according to the Wald  $\chi^2$  values.

The GENMOD output in Table 10.16 contains results in which the standard errors of the estimated coefficients are adjusted for overdispersion. The standard errors are larger by a factor of  $\sqrt{\text{dev}/(N - p)} = \sqrt{4.234}$ . After the adjustment the model terms that appear

**TABLE 10.15 GENMOD Output for Full Model, Wave Solder Data, Example 10.6**

| The GENMOD Procedure                          |              |          |          |            |          |
|---|--------------|----------|----------|------------|----------|
| <i>Model Information</i>                      |              |          |          |            |          |
| Description                                   | Value        |          |          |            |          |
| Data Set                                      | WORK, DEFECT |          |          |            |          |
| Distribution                                  | POISSON      |          |          |            |          |
| Link Function                                 | LOG          |          |          |            |          |
| Dependent Variable                            | Y            |          |          |            |          |
| Observations Used                             | 47           |          |          |            |          |
| <i>Criteria For Assessing Goodness of Fit</i> |              |          |          |            |          |
| Criterion                                     | DF           | Value    | Value/DF |            |          |
| Deviance                                      | 33           | 139.7227 | 4.2340   |            |          |
| Scaled Deviance                               | 33           | 139.7227 | 4.2340   |            |          |
| Pearson Chi-Square                            | 33           | 122.9190 | 3.7248   |            |          |
| Scaled Pearson $\chi^2$                       | 33           | 122.9190 | 3.7248   |            |          |
| Log Likelihood                                | 3837.9339    |          |          |            |          |
| <i>Analysis of Parameter Estimates</i>        |              |          |          |            |          |
| Parameter                                     | DF           | Estimate | Std Err  | Chi Square | Pr > Chi |
| INTERCEPT                                     | 1            | 3.0721   | 0.0344   | 7981.4934  | 0.0001   |
| A   | 1            | 0.1126   | 0.0332   | 11.4710    | 0.0007   |
| B   | 1            | -0.1349  | 0.0344   | 15.3837    | 0.0001   |
| C   | 1            | 0.4168   | 0.0345   | 146.1672   | 0.0001   |
| D   | 1            | -0.0577  | 0.0344   | 2.8118     | 0.0936   |
| E   | 1            | -0.0942  | 0.0334   | 7.9508     | 0.0048   |
| F   | 1            | -0.0176  | 0.0339   | 0.2696     | 0.6036   |
| G   | 1            | -0.3803  | 0.0343   | 122.9944   | 0.0001   |
| AB  | 1            | -0.0175  | 0.0334   | 0.2760     | 0.5993   |
| AC  | 1            | 0.1741   | 0.0339   | 26.3925    | 0.0001   |
| AD  | 1            | 0.0919   | 0.0332   | 7.6421     | 0.0057   |
| BC  | 1            | -0.0788  | 0.0343   | 5.2798     | 0.0216   |
| BC  | 1            | -0.2996  | 0.0344   | 75.8633    | 0.0001   |
| CD  | 1            | 0.0446   | 0.0345   | 1.8275     | 0.1764   |
| SCALE   | 0            | 1.0000   | 0.0000   |            |          |

significant are now *C*, *G*, *AC*, and *BD*. The conclusions concerning the results are quite different when overdispersion is taken into account.

Finally let us consider the reduced model containing the significant model terms that appear in Table 10.16. The fitted model is given by

$$\hat{\mu} = e^{3.0769+0.4405 C-0.4030 G+0.2821 AC-0.3113 BD}$$

where  $\hat{\mu}$  is the estimated mean number of defects. Once again, as in the case of logistic regression, the effects that allow interpretations of the roles of the variables must take into account the multiplicative nature of the model. For example, consider factor *G*, solder temperature. The effect  $e^{-0.4030} = 0.65$  is the *factor* by which the number of defects is reduced when solder temperature is changed from (coded level)  $G = 0$  to  $G = 1$ , the high level. Now consider factor *C*, conveyor speed. The effect of *C* obviously depends on the

**TABLE 10.16    Final Model for Wave Solder Data, Example 10.6**

|   |          |          |          |            |          |
|---|----------|----------|----------|------------|----------|
| Link Function                                 | LOG      |          |          |            |          |
| Dependent Variable                            | Y        |          |          |            |          |
| Observations Used                             | 47       |          |          |            |          |
| <i>Criteria For Assessing Goodness of Fit</i> |          |          |          |            |          |
| Criterion                                     | DF       | Value    | Value/DF |            |          |
| Deviance                                      | 42       | 241.7755 | 5.7566   |            |          |
| Scaled Deviance                               | 42       | 42.0000  | 1.0000   |            |          |
| Pearson Chi-Square                            | 42       | 237.3981 | 5.6523   |            |          |
| Scaled Pearson $\chi^2$                       | 42       | 41.2396  | 0.9819   |            |          |
| Log Likelihood                                | 657.8421 |          |          |            |          |
| <i>Analysis of Parameter Estimates</i>        |          |          |          |            |          |
| Parameter                                     | DF       | Estimate | Std Err  | Chi Square | Pr > Chi |
| INTERCEPT                                     | 1        | 3.0769   | 0.0825   | 1391.0260  | 0.0001   |
| C   | 1        | 0.4405   | 0.0808   | 29.7429    | 0.0001   |
| G   | 1        | -0.4030  | 0.0808   | 24.8954    | 0.0001   |
| AC  | 1        | 0.2821   | 0.0667   | 17.9039    | 0.0001   |
| BD  | 1        | -0.3113  | 0.0808   | 14.8557    | 0.0001   |
| SCALE   | 0        | 2.3993   | 0.0000   |            |          |

NOTE: The scale parameter was estimated by the square root of DEVIANCE/DOF.

level of *A*, the prebake condition. If *A* is at the high level, an increase of conveyor speed from coded level 0 to 1 (one unit) increases the number of defects by a factor of

$$e^{0.4405+0.2821} = e^{0.7226} = 2.05$$

On the other hand, if the prebake condition is at the low level, an increase in *C* from 0 to 1 will increase the number of defects by a factor of 1.17, since

$$e^{0.4405+0.2821} = 1.17$$

Thus the destructive effect of conveyor speed is rendered negligible if the prebake condition is at the low level. The optimum conditions on the factors, i.e., those that result in minimum number of defects, are given by

$$C = -1$$

$$G = +1$$

$$A = +1$$

while factors *B* and *D* are at the extremes, as long as *B* and *D* take opposite signs. The estimated mean number of defects under these conditions is given by  $e^{1.64} = 5.16$ .

### 10.8.8 Diagnostic Plots and Other Aspects of the GLM

As we have previously emphasized, much of the success of RSM centers around graphics. These graphical displays include diagnostic plots of residuals, effect plots, and contour plots. They are described in detail with illustrations, throughout this text, for RSM under standard conditions, that is, normally distributed responses and independent errors with constant variances. In the case of generalized linear models, similar diagnostic plots involving residuals and effects are available, though they require some separate discussion here.

There are two types of residuals that are available in the GLM: the Pearson  $\chi^2$  residual, which has been introduced briefly in Section 10.8.6, and the **deviance residual**. Both the Pearson residual and the deviance residual have as their roots a goodness-of-fit statistic. The Pearson  $\chi^2$  residual is an ordinary residual  $y_i - \mu_i$ , standardized according to the variance  $\hat{v}_i$  at the  $i$ th data point. The sum of squares of these residuals is the **Pearson  $\chi^2$** , which is the basis for a goodness-of-fit statistic in SAS PROC GENMOD. The Pearson  $\chi^2$  is asymptotically  $\chi^2_{N-p}$  like the deviance statistic. However, like the mean deviance, the Pearson  $\chi^2/(N-p)$  is interpreted with the rule of thumb that a value substantially exceeding unity implies that lack of fit is no concern.

In order to gain a sound understanding of the concept of deviance residuals, we should review the basic structure of the deviance calculation with an example. The definition is given in Equation 10.64. Note from this equation that the deviance is a difference in log likelihood. In fact, apart from the factor of 2, we have

$$\lambda(\mathbf{b}) = \ln L(\hat{\boldsymbol{\mu}}) - \ln L(\mathbf{b})$$

where, of course,  $L(\hat{\boldsymbol{\mu}})$  is the likelihood for the saturated model and  $L(\mathbf{b})$  is the likelihood for the model in question. Let us consider an example with a Poisson response. The log likelihood is the log of the joint probability function of  $y_1, y_2, \dots, y_N$ , which is given by the log of

$$P(y_1, y_2, \dots, y_N, \boldsymbol{\mu}) = \frac{e^{-\sum_{i=1}^N \mu_i} \prod_{i=1}^N \mu_i^{y_i}}{\prod_{i=1}^N y_i}$$

Thus

$$\ln L(\hat{\boldsymbol{\mu}}) = - \sum_{i=1}^N \hat{\mu}_i + \sum_{i=1}^N y_i \ln \hat{\mu}_i - \sum_{i=1}^N \ln h_i \quad (10.68)$$

where  $\hat{\mu}_i$  replaces  $\mu_i$  in the log likelihood. For the saturated model  $\hat{\mu}_i$  is replaced by  $y_i$ , the  $i$ th observation, whereas for the Poisson regression model in question,  $\hat{\mu}_i$  is replaced by the

maximum likelihood estimator of the mean from the regression model. As a result, for the Poisson deviance we have

$$\begin{aligned}\lambda(\mathbf{b}) &= 2 \sum_{i=1}^N [(\hat{\mu}_i - y_i) - y_i \ln \hat{\mu}_i + y_i \ln y_i] \\ &= 2 \sum_{i=1}^N [y_i \ln(y_i/\hat{\mu}_i) - (y_i/\hat{\mu}_i)]\end{aligned}\tag{10.69}$$

It can be shown (see Exercise 10.3) that for this case  $\sum_{i=1}^N (y_i - \hat{\mu}_i) = 0$  for the canonical link. However, in general the deviance is of the form  $\sum_{i=1}^N d_i$ , where  $d_i$  is the portion in brackets in Equation 10.69. This  $d_i$  is called the  $i$ th **deviance residual**. It is not exactly what one plots, however. In order to give the residuals the characteristics that more closely resemble those of the residuals in standard linear models, we define the **adjusted deviance residual**  $(r_D)_i$  as

$$(r_D)_i = \text{sgn}(y_i - \hat{\mu}_i) \sqrt{d_i}.$$

This is for most cases the type of residual that is more appropriate to use for plotting purposes. The sum of squares of the  $(r_D)_i$  is the deviance  $\lambda(\beta)$ .

With regard to effect plots, we saw in discussions in our examples that effects are quite often multiplicative in nature, and the interpretations are no less difficult to understand than the linear effects that are generally observed and plotted in the case of linear models. However, these effects often take the form  $e^b$ , where  $b$  is either a regression coefficient or a linear combination of regression coefficients. Since maximum likelihood estimators are known to be asymptotically normal, the best results for normal probability plotting are expected when **standardized coefficients** are plotted, i.e., coefficients divided by their standard errors. It is interesting that when a two-level factorial design or regular fraction is employed, model coefficients are often nearly asymptotically independent when a GLM is used. This depends a great deal on the distribution and link. Indeed, if the so-called **variance-stabilizing link** is used, the coefficients will have covariances that are asymptotically zero. See Myers, Montgomery, Vining, and Robinson (2011). These links include the square root link for the Poisson case, the log link for the exponential distribution, and the log link for the gamma distribution.

**Example 10.7 The Drill Experiment** We consider a  $2^4$  unreplicated factorial that was used to investigate the advance rate of a drill. This response exhibits a nonconstant variance that depends on the level of the response. Daniel (1976) used this experiment to illustrate the usefulness of normal probability plotting. The model fitted to the data here makes use of a log link with an assumed gamma-distributed response. The data appears in Table 10.17.

The use of the log link results in an exponential model. The estimated model coefficients are as follows:



| Term           | Coefficient |
|----------------|-------------|
| Intercept      | 1.598       |
| $x_1$          | 0.065       |
| $x_2$          | 0.290       |
| $x_3$          | 0.577       |
| $x_4$          | 0.163       |
| $x_1x_2$       | -0.017      |
| $x_1x_3$       | 0.005       |
| $x_1x_4$       | 0.034       |
| $x_2x_3$       | -0.025      |
| $x_2x_4$       | -0.008      |
| $x_3x_4$       | 0.049       |
| $x_1x_2x_3$    | 0.005       |
| $x_1x_2x_4$    | 0.025       |
| $x_1x_3x_4$    | 0.027       |
| $x_2x_3x_4$    | -0.017      |
| $x_1x_2x_3x_4$ | 0.019       |

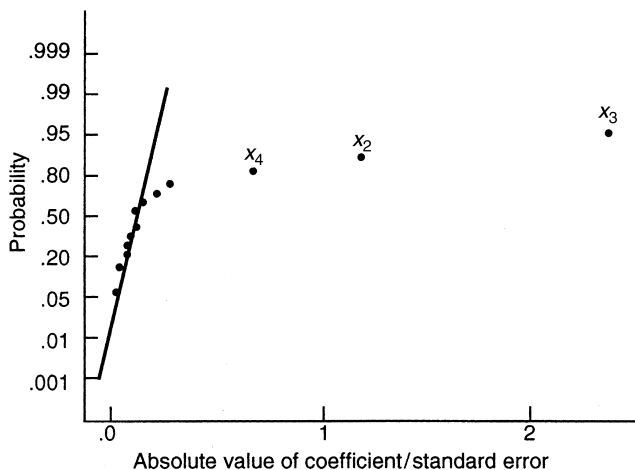
The coefficients divided by their standard errors were plotted on a half normal probability plot. The plot is shown in Fig. 10.14. In this figure  $A = x_1$ ,  $B = x_2$ , and so on.

From the normal probability plot it appears that coefficients for factors  $x_2$ ,  $x_3$ , and  $x_4$  are significantly different from zero. The reduced model was fitted with the resulting estimated response function

$$\hat{y} = e^{1.6032+0.2895x_2+0.5772x_3+0.1656x_4}$$

**TABLE 10.17 The Drill Experiment, Example 10.7**

| Run | $x_1$ | $x_2$ | $x_3$ | $x_4$ | Advance Rate (Original Data) |
|-----|-------|-------|-------|-------|------------------------------|
| 1   | -     | -     | -     | -     | 1.68                         |
| 2   | +     | -     | -     | -     | 1.98                         |
| 3   | -     | +     | -     | -     | 3.28                         |
| 4   | +     | +     | -     | -     | 3.44                         |
| 5   | -     | -     | +     | -     | 4.98                         |
| 6   | +     | -     | +     | -     | 5.70                         |
| 7   | -     | +     | +     | -     | 9.97                         |
| 8   | +     | +     | +     | -     | 9.07                         |
| 9   | -     | -     | -     | +     | 2.07                         |
| 10  | +     | -     | -     | +     | 2.44                         |
| 11  | -     | +     | -     | +     | 4.09                         |
| 12  | +     | +     | -     | +     | 4.53                         |
| 13  | -     | -     | +     | +     | 7.77                         |
| 14  | +     | -     | +     | +     | 9.43                         |
| 15  | -     | +     | +     | +     | 11.75                        |
| 16  | +     | +     | +     | +     | 16.30                        |



**Figure 10.14** Half normal probability plot of effects for the drill data.

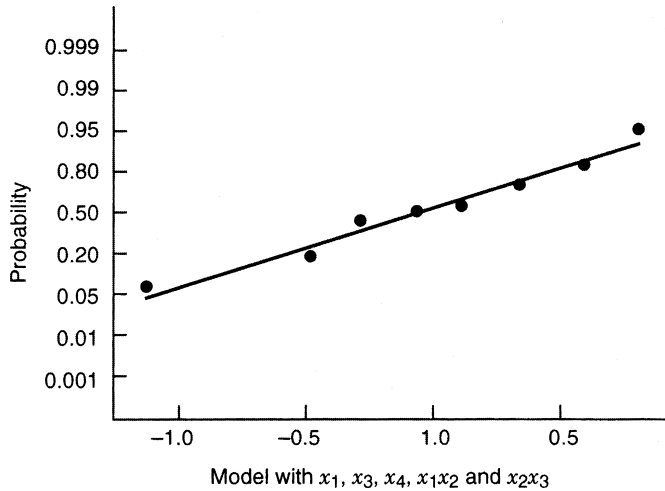
Lewis, Montgomery, and Myers (2001a) conduct a more extensive analysis of this experiment. They illustrate that the use of the GLM is preferable to performing a log transformation and using ordinary least squares, because the GLM produces shorter confidence intervals on the mean response. Lewis, Montgomery, and Myers (2001b) show that this is a reasonable criterion for selecting between a GLM and a standard linear least squares model following a data transformation. Lewis (1998) is also a useful reference.

**Example 10.8 Windshield Slugging** This experiment involves the use of a  $2^{4-1}$  fractional factorial in a study of windshield molding. During the operation debris carried into the die appears as dents, or **slugs**, in the product. Martin, Parker, and Zenick (1987) originally presented and analyzed the data. The purpose of the experiment is to improve slugging performance. The response is the number of good parts out of 1000. The binomial distribution is assumed, and the logit link is used. Table 10.18 gives the data.

The reduced logistic regression model includes the terms  $x_1, x_3, x_4, x_2x_3,$  and  $x_1x_2$ . The reader is asked to verify this in Exercise 10.13. The deviance residuals were plotted for model diagnostic checking and appear in Fig. 10.15. They indicate that the logistic model with the logit link is satisfactory.

**TABLE 10.18** The Windshield Molding Slugging Experiment

| Run | $x_1$ | $x_2$ | $x_3$ | $x_4$ | Good Parts (Original Data) |
|-----|-------|-------|-------|-------|----------------------------|
| 1   | +     | +     | +     | +     | 338                        |
| 2   | +     | +     | -     | -     | 826                        |
| 3   | +     | -     | +     | +     | 350                        |
| 4   | +     | -     | -     | -     | 647                        |
| 5   | -     | +     | +     | -     | 917                        |
| 6   | -     | +     | -     | +     | 977                        |
| 7   | -     | -     | +     | -     | 953                        |
| 8   | -     | -     | -     | +     | 972                        |



**Figure 10.15** Normal probability plots of deviance residuals for logistic model windshield data.

Generalized linear models represent a useful set of techniques for conducting response surface analysis when the response variables are clearly nonnormal and ordinary least squares estimation is not efficient. Modern software packages, including SAS and S-PLUS, have good GLM capability, and JMP can analyze GLMs for the binomial, Poisson, and exponential distributions. Ozol-Godfrey et al. (2008) show how fraction of design space plots, can be adapted to evaluate designs for GLMs. This is a useful tool to add to the kit of a practitioner who wishes to conduct these types of analysis.

## EXERCISES

**10.1** Consider the two simple designs discussed in Section 10.1 namely, the following two-point designs for fitting a first-order model:

- (i)  $N/2$  runs at  $x = -1$ ;  $N/2$  runs at  $x = +1$
- (ii)  $N/2$  runs at  $x = -0.6$ ;  $N/2$  runs at  $x = +0.6$

The region of interest  $[-1, +1]$  in the design variable  $x$ .

- (a) Suppose the first-order model is correct. Compute

$$\frac{N}{2\sigma^2} \int_{-1}^1 \text{Var}[\hat{y}(\mathbf{x})] dx$$

for both designs.

- (b) Suppose the true model is

$$y = \beta_0 + \beta_1 x + \beta_2 x^2$$

with  $\sqrt{N}\beta_2/\sigma = 2.0$ . Compute  $(N/2\sigma^2) \int_{-1}^1 \{\text{Var}[\hat{y}(\mathbf{x})] + [\text{Bias } \hat{y}(\mathbf{x})]^2\} dx$  for both designs.

**10.2** Consider a fitted first-order model in four design variables:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4$$

Suppose the region of interest in the natural variables is given by the following ranges:

$$\xi_1: [100, 150]$$

$$\xi_2: [1, 2]$$

$$\xi_3: [1000, 2000]$$

$$\xi_4: [0, 1]$$

The region is cuboidal. Suppose that the experimenter is interested in protecting against a complete second-order model. A total of 10 runs are to be used.

- (a) Give a minimum bias design in the natural variables.
- (b) Suppose, instead of using the minimum variance design, the analyst achieves protection against model mis-specification by using a scaled second moment of 0.45. Again, two center runs are used in addition to the eight factorial points. Give the appropriate design in the natural variables.

**10.3** Consider Exercise 10.2. Suppose that a first-order model is fitted, but the practitioner is interested in protection against the model

$$E(y) = \beta_0 + \sum_{i=1}^4 \beta_i x_i + \sum_{i < j=2}^4 \sum_{i < j=2}^4 \beta_{ij} x_i x_j$$

In other words, there is no fear of pure quadratic effects in this design region; the protection is sought against *interaction terms only*. Consider the use of Equation 10.18 for the determination of a minimum bias design.

- (a) What are the moment requirements for the minimum bias design in this case?
- (b) Give a minimum bias design with 10 runs for this case (coded levels).
- (c) Give the same design in the natural variables.

**10.4** Refer to Exercise 10.3. With the suggested alternative model, is there any restriction on the second pure design moment for achieving a minimum bias design? Explain.

**10.5** Consider the following general formulation: If the fitted model is first order, and one is interested in protection against a model that contains first-order terms plus two-factor interactions, the minimum bias design is one in which all odd moments through order three are zero. This means that the two-level design for minimum bias must have resolution at least IV. Prove that this statement is correct.

**10.6** Prove the following statement: If the fitted model is first-order in  $k$  variables and one is interested in protecting against a model containing linear terms and two-factor interactions, a minimum mean squared error design is a two-level resolution  $\geq$ IV design with pure second moments (*ii*) as large as possible. As a result, the levels are to be placed at  $\pm 1$  extremes.

**10.7** Reconsider Example 10.2. Show that

$$\text{ASB}_{\text{LS}} = \frac{3\beta_2^2}{\sigma^2} \left( \frac{4l^4}{9} - \frac{4l^2}{9} + \frac{1}{5} \right)$$

and that

$$\text{APV}_{\text{LS}} = 1 + \frac{1}{2l^2}$$

What value of  $l$  results in minimum  $\text{APV}_{\text{LS}}$ ? What is the value of  $\text{ASB}_{\text{LS}}$  at this point? Does this design make sense if you worried about protecting against the quadratic model?

**10.8** Reconsider Example 10.2. Draw a graph of:

- (a) ASB and  $\text{ASB}_{\text{LS}}$  versus  $l$  in the range  $0 \leq l \leq 1$ .
- (b) APV and  $\text{APV}_{\text{LS}}$  versus  $l$  in the range  $0 \leq l \leq 1$ .
- (c) AMSE and  $\text{AMSE}_{\text{LS}}$  versus  $l$  in the range  $0 \leq l \leq 1$ .

Discuss the graphs you have drawn and the potential utility of minimum bias estimation if you are concerned about fitting the wrong model.

**10.9** Consider the same situation described in Example 10.2; now, however, suppose that you wish to use a design with one run at  $\pm l_1$ , and  $\pm l_2$ , with  $l_1 < l_2$ , so that  $N = 4$ .

- (a) Find the minimum bias estimator.
- (b) Find ASB for the minimum bias estimator. Note that this will not be a function of  $l_1$  and  $l_2$ .
- (c) Find APV for the minimum bias estimator as a function of  $l_1$  and  $l_2$ . Prepare a graph that shows how APV changes as  $l_1$  and  $l_2$  change.
- (d) Suppose an experimenter uses the method of least squares with  $l_1 = \pm 0.5$  and  $l_2 = \pm 1$ . Find  $\text{APV}_{\text{LS}}$  and  $\text{PSB}_{\text{LS}}$  for this design. Can you find a design that performs better if you use minimum bias estimation?

**10.10** Consider the case of Poisson response with a log link. Show that  $\sum_{i=1}^N (y_i - \hat{\mu}_i) = 0$ .

**10.11** Table 10.17 shows data from an experiment used to study the advance rate of a drill.

- (a) Verify that the model coefficient estimates are as given in Example 10.7. Use a gamma response with a log link.
- (b) What are the standard errors of the model parameters?
- (c) Find a 95% confidence interval on the mean response at each of the 16 design points. **Hint:** SAS PROC GENMOD will produce these confidence intervals.

**10.12** Analyze the drill experiment in Example 10.7 by ordinary least squares, following a log transformation on the response.

- (a) Does the prediction equation look similar to the one found using the GLM?

- (b) Construct contour plots for the least squares model and the GLM model, with  $x_4 = +1$  and  $x_4 = -1$ , Compare the plots.
- (c) Find a 95% confidence interval on the mean response at each of the 16 design points. Express the interval in terms of the original response.
- (d) Compare the lengths of the confidence intervals from (c) with the lengths of the intervals in Exercise 10.11(c). Which modeling approach would you prefer if response prediction was the objective of the experimenter?
- 10.13** Consider the windshield slugging experiment in Example 10.8. Fit the GLM using a binomial response and the logit link. Verify that the correct model includes the factors  $x_1, x_3, x_4, x_1x_2$ , and  $x_2x_3$ .
- 10.14** The experiment shown in the table below is a  $2^{7-4}$  fractional factorial conducted to study the effect of the seven factors on nonconforming tiles. The response variable is the percentage (out of 1000, say) of tiles that are nonconforming. Analyze this experiment, assuming the binomial distribution and the logit link.

| Run | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | % Nonconforming<br>(Original Data) |
|-----|-------|-------|-------|-------|-------|-------|-------|------------------------------------|
| 1   | +     | +     | +     | +     | +     | +     | +     | 0.16                               |
| 2   | +     | +     | +     | -     | -     | -     | -     | 0.17                               |
| 3   | +     | -     | -     | +     | +     | -     | -     | 0.12                               |
| 4   | +     | -     | -     | -     | -     | +     | +     | 0.06                               |
| 5   | -     | +     | -     | +     | -     | +     | -     | 0.06                               |
| 6   | -     | +     | -     | -     | +     | -     | +     | 0.68                               |
| 7   | -     | -     | +     | +     | -     | -     | +     | 0.42                               |
| 8   | -     | -     | +     | -     | +     | +     | -     | 0.26                               |

- 10.15** Example 2.12 presented the worsted yarn experiment, a  $3^3$  factorial design in which the response variable was the number of cycles to failure of the yarn. A least squares model was fitted to the data after a log transformation.
- (a) Find a GLM for the data, assuming a gamma distribution and the log link.
- (b) Compare the form of this model with the one obtained in Example 2.12. Are they very similar?
- (c) Construct 95% confidence intervals on the mean response at each of the 27 design points. (**Hint:** Use PROC GENMOD to do this for the GLM.) Compare the lengths of the confidence intervals for the two modeling approaches. Based on this comparison, which model would you prefer?
- 10.16** Reconsider the situation described in Exercise 10.15. Use the reciprocal link function, and repeat parts (a) and (b). How does the reciprocal link work in this experiment?
- 10.17** Lewis, Montgomery, and Myers (2001a) present an experiment in a semiconductor manufacturing process involving a CCD where the response variable is the observed number of defects. The data are shown in the following table.

|    | $x_1$  | $x_2$  | $x_3$  | Number of Defects, $y$ | $\sqrt{y}$ |
|----|--------|--------|--------|------------------------|------------|
| 1  | -1     | -1     | -1     | 1                      | 1.000      |
| 2  | 1      | -1     | -1     | 15                     | 3.873      |
| 3  | -1     | 1      | -1     | 24                     | 4.899      |
| 4  | 1      | 1      | -1     | 35                     | 5.916      |
| 5  | -1     | -1     | 1      | 0                      | 0.000      |
| 6  | 1      | -1     | 1      | 13                     | 3.606      |
| 7  | -1     | 1      | 1      | 16                     | 4.000      |
| 8  | 1      | 1      | 1      | 13                     | 3.606      |
| 9  | -1.732 | 0      | 0      | 1                      | 1.000      |
| 10 | 1.732  | 0      | 0      | 17                     | 4.123      |
| 11 | 0      | -1.732 | 0      | 7                      | 2.646      |
| 12 | 0      | 1.732  | 0      | 43                     | 6.557      |
| 13 | 0      | 0      | -1.732 | 14                     | 3.742      |
| 14 | 0      | 0      | 1.732  | 3                      | 1.732      |
| 15 | 0      | 0      | 0      | 4                      | 2.000      |
| 16 | 0      | 0      | 0      | 7                      | 2.646      |
| 17 | 0      | 0      | 0      | 8                      | 2.828      |
| 18 | 0      | 0      | 0      | 6                      | 2.450      |

- (a) Fit a second-order model to the square root of the number of defects. Construct contour plots of the predicted response. The objective is to find a set of operating conditions that minimize the number of defects.
- (b) How well does the least squares model perform as a predictor in the region of interest?
- (c) Repeat part (a) using a Poisson response and the log link.
- (d) Is the GLM in part (c) a better model? Explain.

**10.18** Repeat Exercise 10.17 using the square root link.

**10.19** Reconsider the GLM from the windshield slugging experiment in Example 10.8. Obtain the deviance residuals for the model. Construct a normal probability plot of these residuals. Comment on your findings.

**10.20** Reconsider the GLM for the worsted yarn experiment in Exercise 10.15. Obtain the deviance residuals for the model. Construct a normal probability plot of these residuals. Is there any indication of model inadequacy?

**10.21** Reconsider the GLM for the defects experiment in Exercise 10.16. Obtain the deviance residuals for the model. Construct a normal probability plot of the residuals, and comment on the model's adequacy.

**10.22** Consider the experiment described in Exercise 4.7. Suppose that this experiment had been conducted as a split plot, with factor  $A$  the whole plot variable. Analyze the data, and draw appropriate conclusions.

**10.23** Consider the experiment described in Example 4.2. Suppose that this experiment had been conducted as a split plot, with factor  $B$  the whole plot variable. Analyze the data, and draw appropriate conclusions.

- 10.24** Consider the experiment described in Example 8.5. Suppose that this experiment had been conducted as a split plot, with temperature as the whole plot variable. Analyze the data, and draw appropriate conclusions.
- 10.25** Suppose that you plan to fit a second-order model in three factors. Factor  $A$  is hard to change, but factors  $B$  and  $C$  are easy to change. Suggest an appropriate experimental design for this problem, if the experimenter wants to minimize the total number of runs.
- 10.26** An experimenter wants to fit a second-order model in five factors; factors  $A$  and  $B$  are hard to change, and the other three factors are easy to change.
1. Find a design with the minimum number of runs ( $N = 21$ ).
  2. Find a design with  $N = 32$  runs.
  3. Compare these two designs.
- 10.27** Reconsider the adhesive strength experiment in Example 10.4. What is the minimum number of runs that could be used to investigate the four factors in this problem? Find the design and compare it to the one actually used in Example 10.4.
- 10.28** An experimenter is interested in creating a design involving five factors in the design region  $[-1, +1]^5$  with a total budget of 30 runs.
- (a) Create a sphere packing design.
  - (b) Create a maximum entropy design.
  - (c) Create a Gaussian process  $I$ -optimal design
  - (d) Create a fast flexible filling design.
  - (e) By examining plots, determine which of the following designs:
    - (i) has the smallest average distance between points?
    - (ii) has the largest average distance from the center of the region?
    - (iii) has the best spacing of values for each individual factor across the range  $[-1, +1]$ ?
  - (f) Based on the results in (e), which design do you find most appealing? Why?
- 10.29** Consider a design for three factors with a total of 15 observations.
- (a) Construct an arbitrary Latin hypercube design.
  - (b) Create an algorithm that constructs Latin hypercube designs and that evaluates them based on the *maximin* criterion described in Section 10.4.1. Use the algorithm to find an optimal design based on this criterion.
  - (c) Create an algorithm that constructs Latin hypercube designs and that evaluates them based on the *minimax* criterion (using a grid of points inside the design space) described in Section 10.4.1. Use the algorithm to find an optimal design based on this criterion.
  - (d) Evaluate the best designs found in (b) and (c) for both the minimax and maximin criteria. Which design do you prefer based on their relative efficiencies?
- 10.30** Ahmadi et al. (2005) consider a three-factor face-centered central composite design to explore the effect of reducing the undesirable characteristics of water used in a olive oil production facility. Table E10.1 shows the levels of the scaled factors



**TABLE E10.1** Data for Exercises 10.30, 10.31, and 10.32

| $x_1$ | $x_2$ | $x_3$ | COD<br>Removal | Polyphenol<br>Removal | Color<br>Removal | Aromatocity<br>Removal |
|-------|-------|-------|----------------|-----------------------|------------------|------------------------|
| -1    | -1    | -1    | 3.20           | 7.80                  | 1.33             | 2.15                   |
| 1     | -1    | -1    | 41.08          | 99.90                 | 17.16            | 9.64                   |
| -1    | 1     | -1    | 1.85           | 4.50                  | 0.69             | 0.91                   |
| 1     | 1     | -1    | 21.26          | 51.80                 | 8.88             | 4.08                   |
| -1    | -1    | 1     | 9.89           | 21.20                 | 1.38             | 8.33                   |
| 1     | -1    | 1     | 43.94          | 94.13                 | 17.72            | 37.44                  |
| -1    | 1     | 1     | 4.10           | 8.70                  | 0.72             | 3.44                   |
| 1     | 1     | 1     | 18.72          | 40.10                 | 9.20             | 15.45                  |
| -1    | 0     | 0     | 30.23          | 52.35                 | 18.50            | 17.52                  |
| 1     | 0     | 0     | 55.76          | 99.50                 | 34.80            | 32.78                  |
| 0     | -1    | 0     | 23.14          | 54.50                 | 11.67            | 9.52                   |
| 0     | 1     | 0     | 12.15          | 30.14                 | 6.10             | 5.50                   |
| 0     | 0     | -1    | 30.24          | 70.42                 | 12.15            | 6.69                   |
| 0     | 0     | 1     | 27.55          | 61.32                 | 9.12             | 12.20                  |
| 0     | 0     | 0     | 38.60          | 71.10                 | 19.48            | 17.86                  |

and the results of the experiment on the four responses. The three factors that were considered in the experiment were the  $\text{H}_2\text{O}_2$ -to-Fe(II) ratio (levels 1.67, 5, 8.33), pH (levels: 3,4,5) and olive oil mill wastewater (OMW concentration levels: 40%, 70%, 100%). The responses represent the percentage of the characteristic removed, and hence large values are desirable as a sign of improved restoration of the wastewater to pristine condition.

- Fit a second-order model to the Polyphenol response based on the three factors.
- By examining a contour plot of the estimated response throughout the design region, where are there problems? Explain why this would be problematic for the experimenters.
- Using the logit link of Equation 10.58, fit a logistic regression model to the Polyphenol response.
- By examining a contour plot of the estimated response in the design region, compare the results to those obtained in (b). Does this model correct the problem found previously? Explain.

**10.31** Reconsider the experiment in Exercise 10.30 and Table E10.1.

- Fit a logistic regression model to the response Aromatocity removal. What is the estimated maximum percentage of aromatocity that can be removed from factor combinations within the design space?
- Where in the design space is the optimum achieved?
- Using the results from Exercise 10.30(d), what is the maximum estimated percentage of Polyphenol that can be removed? Where is this optimum achieved?
- Find the maximum estimated amount of Aromatocity that can be removed, subject to the amount of Polyphenol removed being at least 95%.

- 10.32** Reconsider the experiment in Exercise 10.30 and 10.31, with data in Table E10.1.
- (a) Fit a logistic regression model to the response Color removal. What is the estimated maximum percentage of color that can be removed from factor combinations within the design space? Where in the design space is the optimum achieved?
  - (b) Based on the results from Exercise 10.31 and (a), which responses have greater trade-offs when trying to optimize the amount removed?
  - (c) Construct a Pareto front based on optimizing both color and aromaticity removal. What is the range of values for each of responses found on the Pareto front? Which location in the design space are involved in solutions found on the Pareto front?
  - (d) Using an additive desirability function with individual desirability scores based on the range of values found on the Pareto front, which locations in the design space are best for equal weighting of color and aromaticity?
  - (e) Repeat (d) using individual desirability scores based on the range of values for each response found across the entire design space. How do the optimal locations compare for the two scalings? Which scaling seems preferable? Explain.

### Short Answer Questions

- 10.33** Answer the following short answer questions:
- (a) Justify how using a mean squared error criterion for evaluating the goodness of a design might be beneficial.
  - (b) Describe the criteria for creating a minimum bias design for a first-order model.
  - (c) Describe how difficulty in setting the input factors can impact the results of the design and analysis of an experiment.
  - (d) Justify why it is important to consider different types of designed experiments when the response will be obtained from a computer model. How are the data obtained from these types of experiments different?
  - (e) Explain the difference between stochastic and deterministic simulation models.
  - (f) Explain why it is important to use a metric like minimax or maximin to evaluate different designs within a class of designs (like Latin hypercube designs)?
  - (g) How does the curse of dimensionality with increasing numbers of input factors influence the choice of designed experiments?
  - (h) Give a situation when it would be important to use a split-plot design rather than a completely randomized design.
  - (i) Describe how the construction of a good split-plot design becomes more complicated than for a completely randomized design.
  - (j) Explain why different designs are needed when considering a generalized linear model.
  - (k) Describe the role of the link function in a generalized linear model. Why does it make a difference which link is chosen?

**10.34** Answer the following True/False questions:

- (a) Minimum bias designs are frequently the same as  $D$ -optimal designs.
- (b) Designs for computer experiments and physical experiments typically use the same design criteria for optimization.
- (c) The criteria for split-plot designs need to be adapted to take into account the different error structure of the model.
- (d) Model bias can be estimated regardless of the choice of design.
- (e) In general, it is better to use a higher-degree model to construct an optimal design than to use too simple of a model.
- (f) Minimax designs use only the design points to estimate the criterion value, while maximin designs use the design space and the design points.
- (g) All Latin hypercube designs for a particular design set-up will perform similarly for estimation of the model.
- (h) Using replicates for a computer experiment only makes sense if the response is stochastic, not deterministic.

**10.35** Circle the correct answer for each question:

- (a) For a split-plot design
  - (i) Standard least squares estimation is the appropriate method to estimate the model parameters.
  - (ii) The choice of best design is dependent on the relative size of the two error terms.
  - (iii) The designation of hard-to-change and easy-to-change factors is randomized.
  - (iv) None of (i)–(ii) are correct statements.
  - (v) All of (i)–(iii) are correct statements.
- (b) For a generalized linear model with binomial data,
  - (i) The canonical link is the logit link.
  - (ii) The variance of the response is dependent on its mean.
  - (iii) Estimation of model parameters is typically done with the method of maximum likelihood.
  - (iv) None of (i)–(iii) are correct statements.
  - (v) All of (i)–(iii) are correct statements.



---

# 11

---

## ROBUST PARAMETER DESIGN AND PROCESS ROBUSTNESS STUDIES

### 11.1 INTRODUCTION

In the 1980s, Genichi Taguchi [Taguchi and Wu (1980), Taguchi (1986, 1987)] introduced new ideas on quality improvement in the United States. He proposed an approach he called **parameter design** for reducing variation in products and processes. His methods and philosophy generated considerable interest among quality engineers and statisticians. During the 1980s his methodology gained usage at AT&T Bell Laboratories, Ford Motor Company, Xerox, and many other prominent organizations. These techniques generated much debate and controversy in the statistical community—not about Taguchi’s philosophy, but rather about its implementation and the technical nature of data analysis. As a result, a response surface approach evolved and has become a collection of tools that allow one to adopt Taguchi’s robust design concept while providing a more statistically sound and efficient approach to analysis. This RSM approach to parameter design is the subject of this chapter.

### 11.2 WHAT IS PARAMETER DESIGN?

Parameter design, or **robust parameter design** (RPD), is essentially a principle that emphasizes proper choice of levels of controllable factors (**parameters** in Taguchi’s terminology) in a system. The system can be a process or a product. The principle of choice of levels focuses to a great extent on variability around a prechosen target for the output response. These controllable factors are called **control factors**. It is assumed that the majority of variability around target is caused by the presence of a second set of factors called noise

---

*Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, Fourth Edition.  
by Raymond H. Myers, Douglas C. Montgomery, and Christine M. Anderson-Cook.  
Copyright © 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

factors or **noise variables**. Noise factors are **uncontrollable** in the design of the product or in the normal operation of the process. In fact, it is this lack of control that transmits variability to the process response. As a result, the term **robust parameter design** entails designing (not in the sense of experimental design) the system (selecting the levels of the controllable variables) so as to achieve **robustness** (insensitively) to inevitable changes in the noise variables. Often an RPD study in a manufacturing process is called a **process robustness study**.

A significant difference between Taguchi's view of process optimization and the traditional viewpoint centers around the notion of *reducing process variability* by involving factors that cause process variability in the experimental design, modeling, and optimization exercise. These factors, the noise factors, are often functions of environmental conditions—for example, humidity, properties of raw materials, and product aging. In certain applications they may involve the way the consumer handles or uses the product. Noise variables may be, and often are, controlled at the research or development level, but of course they cannot be controlled at the production or product use level.

### 11.2.1 Examples of Noise Variables

It should not be inferred that Taguchi discovered the notion of noise variables. Prior to Taguchi, researchers certainly were aware that certain uncontrollable factors provide major sources of variability. However, Taguchi encouraged the formal use of noise variables in the experimental design and, as a result, allowed the analysis to involve process variability. Table 11.1 presents some examples of noise variables and control variables in several scenarios.

Table 11.1 gives only a few examples where the noise variables involved clearly cannot be controlled in the process, but could be controlled in an experimental situation—that

**TABLE 11.1** Some Examples of Control Variables and Noise Variables

| Application   | Control Variables  | Noise Variables  |
|---|--|--|
| Development of a cake mix   | Amount of sugar, flour, and other dry ingredients                            | Oven temperature, baking time, amount of milk added                                      |
| Development of a gasoline   | Ingredients in the blend; other processing conditions                        | Type of driver, driving conditions, engine type  |
| Operation of a tablet press to manufacture a pharmaceutical product | Active ingredient types and concentrations; binder material production speed | Relative humidity, temperature   |
| Large-scale chemical process  | Processing conditions, including nominal ambient temperature                 | Deviations from nominal ambient temperature; deviations from other processing conditions |
| Production of a box-filling machine for filling boxes of detergent  | Surface area; geometry of the machine (rectangular, circular)                | Particle size of detergent   |
| Manufacturing a dry detergent                                       | Chemical formulation, processing variables                                   | Temperature and relative humidity during manufacture                                     |

is, in a designed experiment. One important point should be made about the “large-scale chemical process” example. Many times the most important noise variables are tolerances in the control variables. In the chemical and some other industries, so-called controllable variables are not controlled at all. In fact, if there is a nominal temperature of, say, 250°F during the large-scale production process, the temperature may vary between 235 and 265°F, and the inability to control the process temperature produces unwanted variability in the product. As a result, nominal temperature (or nominal pressure) becomes a control variable, and the tolerance on temperature at levels  $\pm 15^\circ\text{F}$  (say) is a noise variable.

### 11.2.2 An Example of Robust Product Design

Consider a situation in which one is interested in establishing the geometry and surface area of a box-filling machine that achieves a rate of filling that is “closest” to 14 g/sec. We will call surface area *factor A* (two levels) and geometry *factor B* (two levels). The single noise variable is the particle size of the detergent product. Three levels of the noise variable were used. Suppose that we can sort the detergent by particle size into the three levels. The data are as follows:

| Particle Size 1 |      |      | Particle Size 2 |      |      | Particle Size 3 |      |      |
|-----------------|------|------|-----------------|------|------|-----------------|------|------|
| <i>B</i>        |      |      | <i>B</i>        |      |      | <i>B</i>        |      |      |
| <i>A</i>        | -1   | +1   | <i>A</i>        | -1   | +1   | <i>A</i>        | -1   | +1   |
| -1              | 13.7 | 13.7 | -1              | 14.9 | 14.2 | -1              | 17.4 | 14.4 |
| +1              | 14.0 | 11.9 | +1              | 16.0 | 11.8 | +1              | 12.0 | 11.7 |

Clearly the product that is most insensitive to the noise factor, particle size, is given by (+1, +1). The sample standard deviation of fill rate is  $s = 0.1\text{g/sec}$ . However, the mean filling rate across particle size is 11.8, and the target is 14.0. On the other hand, the product (+1, -1) results in an average of 14.0, but the variability around the target is relatively large ( $s = 2.0$ ) and thus the product is *not a robust product*. Consider the product (-1, +1). The average filling rate is 14.1, and the variability around the target is quite small ( $s = 0.36$ ). Thus, the optimum design among the four is (-1, +1). It produces the best combination of bias error and variance error.

Strictly speaking, the most robust parameter design is the one that is insensitive to changes in the noise variables. The implication of this definition is that the robust product minimizes variance transmitted by the noise variables. However, one cannot ignore the process mean. In this illustration, one might consider the squared error loss criterion; that is, choose the product that minimizes

$$L = \sum (y - 14)^2$$

where the summation is taken over the levels of the particle size. In later sections where we deal with the response surface approach to robust parameter design and process robustness studies, the process mean and variance will be considered simultaneously.

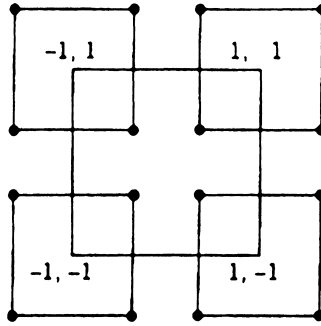


Figure 11.1 The  $2^2 \times 2^2$  crossed array.

### 11.3 THE TAGUCHI APPROACH

#### 11.3.1 Crossed Array Designs and Signal-to-Noise Ratios

Taguchi’s methodology for the RPD problem revolves around the use of orthogonal designs where an **orthogonal array** involving control variables is crossed with an orthogonal array for the noise variables. For example, in a  $2^2 \times 2^2$ , the  $2^2$  for the control variables is called the **inner array** and the  $2^2$  for the noise variables is called the **outer array**. The result is a 16-run design called the **crossed array**. Figure 11.1 gives a graphical representation.

The corners of the inner array are represented by  $(-1, -1)$ ,  $(-1, +1)$ ,  $(+1, -1)$ , and  $(+1, +1)$  for the control variables. Each outer array is a  $2^2$  factorial in the noise variables. The dots in the outer arrays represent the locations of observations. Figure 11.2 shows another crossed array with 32 observations. The design is a  $2^{4-1}$  as the inner array, involving the control variables,  $x$ ’s, crossed with a  $2^{3-1}$  as the outer array, involving the noise variables,  $z$ ’s.

Taguchi suggested that one may summarize the observations in each outer array with a summary statistic that provides information about the mean and variance. For example, in Figs. 11.1 and 11.2, the summary statistic is computed across four observations. Thus, in Fig. 11.2 there will be eight summary statistics. The summary statistic is called a **signal-to-noise ratio (SNR)**, and the statistical analysis is done using the SNR as the response

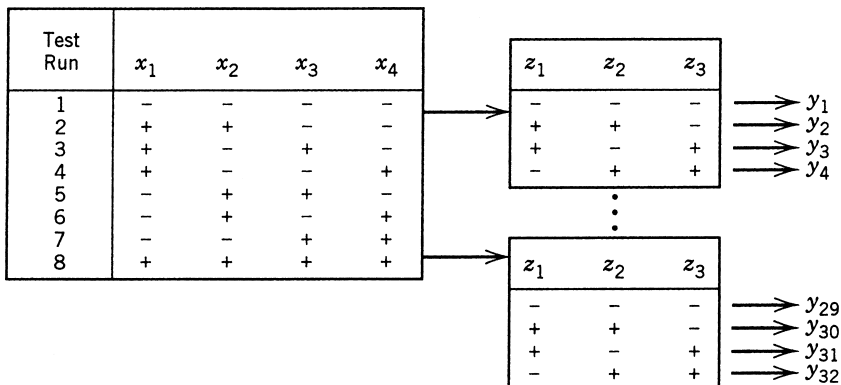


Figure 11.2 The  $2^{4-1} \times 2^{3-1}$  crossed array.



variable. There are many different SNRs. However, there are four primary ones suggested by Taguchi. The choice depends on the goal of the experiment. As in the case of response surface work, there are three specific goals:

1. *The smaller the better.* The experimenter wishes to minimize the response.
2. *The larger the better.* The experimenter wishes to maximize the response.
3. *The target is best.* The experimenter wishes to achieve a particular target value.

**The Smaller the Better** Taguchi treats the smaller-the-better case as if there were a target value of zero for the response. Thus, the quadratic-loss function  $E_z(y - 0)^2$  leads to the performance criterion

$$\text{SNR}_s = -10 \log \sum_{i=1}^n \frac{y_i^2}{n} \quad (11.1)$$

The rationale comes from the need to find  $\mathbf{x}$ , whose elements are the levels of the control variables that minimize the expected squared error  $E_z(y^2)$ , where  $E_z$  implies expectation across the distribution of the noise variables  $\mathbf{z}$ . In the SNR calculation,  $\sum_{i=1}^n$  implies summation over  $n$  response values at the outer array points. Thus, an SNR value will be present at each inner array design point. Because the  $-10 \log$  transformation is used, one seeks to maximize SNRs.

**The Larger the Better** This case is treated in the same fashion as the smaller-the-better case, but  $y_i$  in Equation 11.1 is replaced by  $1/y_i$ . Thus, the SNR is motivated by the squared error criterion  $E_z(1/y)^2$ . As a result, the SNR is given by

$$\text{SNR}_l = -10 \log \sum_{i=1}^n \frac{1/y_i^2}{n} \quad (11.2)$$

Again, values of the control variables are sought that maximize SNR.

**The Target is Best** In this case, we are attempting to determine values of  $\mathbf{x}$  that achieve a target value for the response, namely  $y = t$ . Deviations in either direction are undesirable. Two different SNRs are considered. The appropriate one depends on the nature of the system. If the response mean and variance can be altered independently, Taguchi suggests that one or more **tuning factors** be used to eliminate bias—that is, provide an adjustment that results in  $E_z(y) = t$ . These tuning factors allow the analyst to alter the mean but leave the variance unchanged. As a result, the analysis is a two-step process: Select the tuning factors that render  $y = t$ , and then select levels of other control factors that maximize an SNR. Assuming that this can be accomplished, the obvious squared error loss function,  $E_z(y - t)^2$ , reduces to  $\text{Var}(y)$ . Thus, the SNR used by Taguchi is given by

$$\text{SNR}_{T_1} = -10 \log s^2 \quad (11.3)$$

where  $s^2 = \sum_{i=1}^n (y_i - \bar{y})^2 / (n - 1)$ . Thus,  $s^2$  is the sample variance over the outer array design points. Again, values of  $\mathbf{x}$  are chosen so that SNR is maximized.

A SNR is suggested by Taguchi for cases in which the response standard deviation is related to the mean. It nicely accommodates situations in which the relationship is linear. In this case adjustment or tuning factors are sought that, again, eliminate bias but leave the coefficient of variation  $\sigma_y/\mu_y$  relatively unaffected. As a result, the natural SNR involves the sample coefficient of variation  $s/\bar{y}$ . In fact, Taguchi's SNR for this case is given by

$$\text{SNR}_{T_2} = -10 \log \left( \frac{\bar{y}^2}{s^2} \right) \quad (11.4)$$

where the SNR is to be maximized. So, again, the two-step procedure is used. Values of the tuning factor (or factors) are chosen that render  $y = t$ . These tuning factors have an effect on the mean response but little or none on the SNR.

**What Is the Utility of the SNR?** The purpose of the SNR in Taguchi's approach to robust parameter design is to provide an easy-to-use performance criterion that takes the process mean and variance into account. The terminology itself is open to criticism. Of the four major SNR formulations, only the one in Equation 11.4 involves a true signal-to-noise ratio (because it is the only one that is dimensionless). There is a wealth of literature in which SNRs are criticized. See, for example, Box (1988) and Welch et al. (1990). A panel discussion, edited by Nair (1992), provides opinions by several researchers regarding the use of SNR. We would like to focus on one interesting characteristic of the SNRs here. While the intention is to use them as performance characteristics that take both response mean and variance into account, many authors have suggested the use of *separate models* for the process mean and variance as a way of achieving better understanding of the process itself. More details will be provided on this approach in a later section.

The use of a SNR does not guarantee that the subject matter scientist will secure valuable information regarding the roles of the process mean and process variance. This **uncoupling** of those control factors that affect mean (location effects) and control factors that affect variance (dispersion effects) is very important in the understanding of the process. For example, the SNR in Equation 11.4 can be written

$$\text{SNR}_{T_2} = -10 \log \bar{y}^2 + 10 \log s^2$$

So, of course, the maximization of this SNR does not distinguish which control factors are location effects and which are dispersion effects. As another example, consider  $\text{SNR}_s$  in Equation 11.1

$$\text{SNR}_s = -10 \log \left( \sum y_i^2 / n \right)$$

While  $\sum y_i^2 / n$  does deal with variability around the target of zero, it is clear that an analysis using this SNR cannot separate the location effects from dispersion effects. It is easy to see that

$$\frac{\sum y_i^2}{n} = \bar{y}^2 + \left( 1 - \frac{1}{n} \right) s^2$$

so again, the mean and variance contributions in  $\text{SNR}_s$  are confounded.

Because the SNRs generally confound location and dispersion, they are not necessarily unique. That is, there can be many different samples of data that will result in exactly the same value of the SNR. Thus, it is not necessarily clear how one should interpret a particular value of a SNR.

**Crossed Array Designs** A final point should be made about crossed array designs. When they were executed it seems that many of these designs were actually run as split plots. That is, a run (or treatment combination) in the inner array was fixed, and all outer array runs were conducted at that setting of the control factors. Another possibility is to fix the settings of the noise variables at a particular run in the outer array and then conduct all runs in the inner array at that configuration of the noise variables. Either approach will produce a split-plot structure. The latter approach is generally preferable, because the control factors and their interactions with the noise factors would be more precisely estimated as they reside in the subplot. For more discussion, see Box and Jones (1992, 2000). No split-plot structure is considered in Taguchi's analysis, which we now describe.

### 11.3.2 Analysis Methods

Taguchi's notion of quality improvement places emphasis on variance reduction. His approach to variance reduction is widely acknowledged as a very important contribution to statistics and engineering. While others had suggested the use of noise variables before Taguchi, his idea of linking their use to variance reduction is original and has had a profound impact. Details of his analysis ideas can be found in Taguchi (1987).

To understand the rationale behind the Taguchi analysis, one must be reminded that the purpose of parameter design is to find the settings on the control variables that produce a robust product or process. The control variables may be quantitative or qualitative. As a result, the Taguchi analysis reduces to process optimization when the performance criterion is the SNR.

The analysis method is simple and is very much linked to the crossed array concept. There is an attempt to **maximize** the SNR. The modeling that is done essentially relates the SNR to the control variables in a main-effects-only scenario. Taguchi rarely considers interactions among the control variables. In fact, many of the designs advocated by Taguchi do not allow estimation of these interactions. Standard ANOVA techniques are used to identify the control factors that affect the SNR. In addition, ANOVA is done on the average response  $\bar{y}$  in order to ascertain possible adjustment factors. The latter are set to levels that bring the mean to target (in the target-is-best case), while the factors that affect the SNR are set to levels that maximize it. This completes the two-step process where it is appropriate. The setting of levels is accomplished by a **pick-the-winner** analysis in which a marginal means approach is accomplished. To gain further insight, consider the following example.

**Example 11.1 Wave Soldering Optimization** In an experiment described by Schmidt and Launsby (1990), solder process optimization is accomplished by robust parameter design in a printed circuit board assembly plant. After the components are inserted into a bare board, the board is put through a wave solder machine, which mechanically and electrically connects all the components into the circuit. Boards are placed on a conveyor and then put through the following steps: bathing in a flux mixture to remove oxide, preheating to minimize warpage, and soldering. An experiment is designed to determine

the conditions that give minimum numbers of solder defects per million joints. The control factors and levels are as follows:

| Factor                      | (-1) | (+1) |
|-----------------------------|------|------|
| A, solder temperature (°F)  | 480  | 510  |
| B, conveyor speed (ft/min)  | 7.2  | 10   |
| C, flux density             | 0.9  | 1.0  |
| D, preheat temperature (°F) | 150  | 200  |
| E, wave height (in.)        | 0.5  | 0.6  |

In this illustration, three noise factors are difficult to control in the process. These are the solder temperature, the conveyor speed, and the assembly type. Often in such processes, the natural noise factors are *tolerances in the control variables*. In other words, as these factors vary off their nominal values, variability is transmitted to the response. It is known that the control of temperature is within  $\pm 5^\circ\text{F}$  and that the control of conveyor speed is within  $\pm 0.2$  ft/min. Thus, it is conceivable that variability is increased substantially because of the inability to control these two factors at nominal levels. The third noise factor is the assembly type. One of the two types of assembly will be used. Thus, the noise factors are as follows:

| Factor                     | (+1) | (-1) |
|----------------------------|------|------|
| F, solder temp. (°F)       | 5    | -5   |
| G, conveyor speed (ft/min) | +0.2 | -0.2 |
| H, assembly type           | 1    | 2    |

Both the control array (inner array) and the noise array (outer array) were chosen to be fractional factorials. The inner array is a  $2^{5-2}$  design, and the outer array is a  $2^{3-1}$ . The *crossed array* and the response values are shown in Table 11.2.

TABLE 11.2 The Wave Solder Experiment

| Inner Array |    |    |    |    | Outer Array |    |    |    | (SNR) <sub>s</sub> |        |
|-------------|----|----|----|----|-------------|----|----|----|--------------------|--------|
| A           | B  | C  | D  | E  | F           | G  | H  |    |                    |        |
| 1           | 1  | 1  | -1 | -1 | -1          | 1  | 1  | -1 | 275                | -46.75 |
| 1           | 1  | -1 | 1  | 1  | -1          | 1  | -1 | 1  | 136                | -42.61 |
| 1           | -1 | 1  | -1 | 1  | -1          | -1 | 1  | 1  | 185                | -47.81 |
| 1           | -1 | -1 | 1  | -1 | -1          | 1  | -1 | 1  | 47                 | -39.51 |
| -1          | 1  | 1  | 1  | -1 | 1           | -1 | 1  | -1 | 295                | -48.15 |
| -1          | 1  | -1 | -1 | 1  | 1           | 1  | -1 | -1 | 234                | -45.97 |
| -1          | -1 | 1  | 1  | 1  | 1           | -1 | -1 | -1 | 328                | -45.76 |
| -1          | -1 | -1 | -1 | -1 | 1           | 1  | 1  | -1 | 186                | -43.59 |

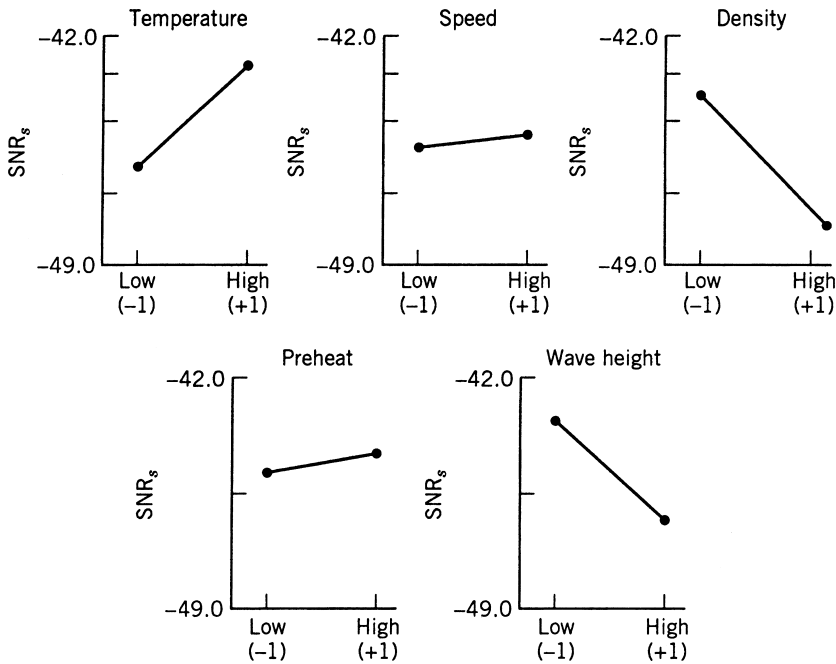


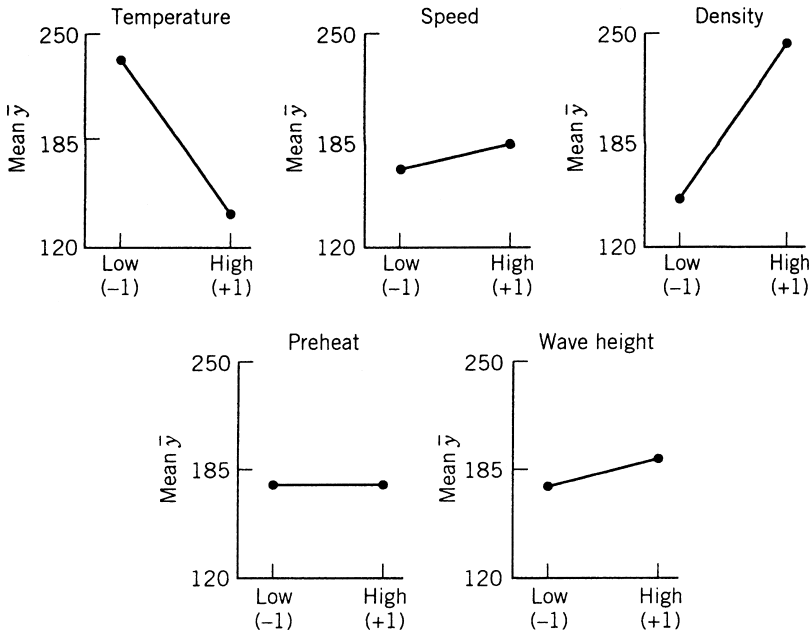
Figure 11.3 Plot of  $SNR_s$  against level for each control factor.

**Analysis** The analysis in this example involves a marginal means modeling of the  $SNR$  of Equation 11.1, namely  $SNR_s$ , because it is of interest to minimize the number of solder defects. One analytical device used frequently by Taguchi is to depict a main-effects-only analysis through the graphs in Fig. 11.3. The mean  $SNR_s$  is plotted against levels of each control factor. The means are taken across levels of the other factors. In addition to the mean  $SNR_s$  plot, a plot of  $\bar{y}$  against the levels of the control variables is also made and appears in Fig. 11.4.

It is clear from Figs. 11.3 and 11.4 that temperature and flux density are critical control factors that have a profound effect on  $SNR_s$  and  $\bar{y}$ . It also appears that wave height has an effect on  $SNR_s$  but little or no effect on  $\bar{y}$ . Because it is of interest to maximize  $SNR_s$  and minimize  $\bar{y}$ , the following represent suggested operating conditions:

$$\begin{aligned}
 \text{Solder temperature} &= 510^\circ\text{F} \\
 \text{Flux density} &= 0.9 \\
 \text{Wave height} &= 0.5 \text{ in.}
 \end{aligned}
 \tag{11.5}$$

The conveyor speed and preheat temperature can be placed at the most economical settings, presumably at low levels. The effect of wave height is marginal compared to the impact of solder temperature and flux density. The analysis suggests that the conditions given by Equation 11.5 are the most **robust conditions**, that is, in the context of this example, those conditions that are most insensitive to changes in the noise variables.



**Figure 11.4** A plot of  $\bar{y}$  against the levels of control factors.

**Further Comments on the Analysis** In addition to earlier concerns regarding the appropriateness of the SNR, other comments concerning the analysis are in order. The plots in Figs. 11.3 and 11.4 are appropriate for selecting optimum (most robust) conditions *if one is willing to assume no interaction among the control variables*. Obviously, in many areas of application this is a safe assumption, and certainly Taguchi's success bears this out. However, there are many areas in which interactions have a very large contribution to process or product performance. This renders the plots of main effect means highly misleading. For example, in Fig. 11.3, consider the temperature plot. High positive slope reveals that across levels of the other factors an increase in temperature increases SNRs. But suppose when the density is low (which is an important level to maintain) the effect of temperature is negligible or even negative. Then, there will be little need to use high temperature when one operates with low density.

Many of the designs employed by Taguchi for inner arrays, where control factors reside, do not accommodate interactions. In addition to the apparent difficulty with SNRs, this has led to criticism of the methodology suggested by Taguchi. Many of the designs suggested by Taguchi are saturated or near-saturated Plackett–Burman designs and thus do not allow the experimenter to estimate interaction effects among the control variables.

A major component of the methodology of Taguchi's robust design is to use **confirmatory trials** at the recommended operating conditions. Indeed, this recommendation is quite appropriate for any statistical procedure in which optimum operating conditions are estimated.

**Example 11.2 The Connector Pulloff Force Experiment** An experiment was conducted in order that a method could be found to assemble an elastometric connector to a nylon tube that would deliver the required pulloff information for use in an automotive

**TABLE 11.3 Factors and Levels on the Connector Pulloff Force Experiment**

|   | Levels  |        |       |
|---|---------|--------|-------|
| <i>Controllable Factors</i>               |         |        |       |
| A = Interference                          | Low     | Medium | High  |
| B = Connector wall thickness              | Thin    | Medium | Thick |
| C = Insertion depth                       | Shallow | Medium | Deep  |
| D = Percent adhesive in connector pre-dip | Low     | Medium | High  |
| <i>Uncontrollable Factors</i>             |         |        |       |
| E = Conditioning time                     | 24 hr   | 120 hr |       |
| F = Conditioning temperature              | 72°F    | 150°F  |       |
| G = Conditioning relative humidity        | 25%     | 75%    |       |

engine application. The objective is to maximize the pulloff force [see Byrne and Taguchi (1987)]. Four variables were used as control variables. They are interference (*A*), connector wall thickness (*B*), insertion depth (*C*), and percent adhesive in connector pre-dip (*D*). Three factors were chosen as noise variables because they are difficult to control in the process and make a major contribution to variability in the pulloff force. They are conditioning time (*E*), conditioning temperature (*F*), and conditioning relative humidity (*G*). The levels are shown in Table 11.3.

Note that the control factors are at three levels, while the noise factors are at two levels. The experimental design and measured pulloff force are given in Table 11.4. Notice that the inner array is a  $3^{4-2}$  fractional factorial design. (Taguchi refers to this as an  $L_9$ .) The outer array in the noise variables is a  $2^3$ . The levels for the control variables are  $-1, 0,$  and  $+1$ —representing low, medium, and high, respectively. The usual  $-1, +1$  notation is used for low and high as the noise variables. Note also that the design is a crossed array, giving a total of  $9 \times 8 = 72$  observations. The  $SNR_i$  and  $\bar{y}$  values are given. If formal modeling were to be done, then the fraction of the  $3^4$  could accommodate linear and quadratic terms in each control variable. However, *there are no degrees of freedom left for interaction among the control variables*. The analysis pursued is, once again, the marginal means plots for  $SNR_i$  and  $\bar{y}$ . Figures 11.5 and 11.6 present the marginal means plots for  $\bar{y}$  and  $SNR_i$ , respectively.

**TABLE 11.4 Inner and Outer Array for Example 11.2**

| Run | Inner Array ( $L_9$ ) |          |          |          | Outer Array ( $L_8$ ) |          |          |          |          |          |          |          | Responses |        |          |           |         |      |      |      |      |
|-----|-----------------------|----------|----------|----------|-----------------------|----------|----------|----------|----------|----------|----------|----------|-----------|--------|----------|-----------|---------|------|------|------|------|
|     |                       |          |          |          | <i>E</i>              | <i>F</i> | <i>G</i> | <i>E</i> | <i>F</i> | <i>G</i> | <i>E</i> | <i>F</i> |           |        | <i>G</i> | $\bar{y}$ | $SNR_i$ |      |      |      |      |
|     | <i>A</i>              | <i>B</i> | <i>C</i> | <i>D</i> | $-1$                  | $-1$     | $-1$     | $-1$     | $+1$     | $+1$     | $+1$     | $-1$     | $-1$      | $-1$   | $+1$     | $+1$      | $+1$    | $-1$ | $-1$ | $-1$ | $+1$ |
| 1   | -1                    | -1       | -1       | -1       | 15.6                  | 9.5      | 16.9     | 19.9     | 19.6     | 19.6     | 20.0     | 19.1     | 17.525    | 24.025 |          |           |         |      |      |      |      |
| 2   | -1                    | 0        | 0        | 0        | 15.0                  | 16.2     | 19.4     | 19.2     | 19.7     | 19.8     | 24.2     | 21.9     | 19.475    | 25.522 |          |           |         |      |      |      |      |
| 3   | -1                    | +1       | +1       | +1       | 16.3                  | 16.7     | 19.1     | 15.6     | 22.6     | 18.2     | 23.3     | 20.4     | 19.025    | 25.335 |          |           |         |      |      |      |      |
| 4   | 0                     | -1       | 0        | +1       | 18.3                  | 17.4     | 18.9     | 18.6     | 21.0     | 18.9     | 23.2     | 24.7     | 20.125    | 25.904 |          |           |         |      |      |      |      |
| 5   | 0                     | 0        | +1       | 1        | 19.7                  | 18.6     | 19.4     | 25.1     | 25.6     | 21.4     | 27.5     | 25.3     | 22.825    | 26.908 |          |           |         |      |      |      |      |
| 6   | 0                     | +1       | -1       | 0        | 16.2                  | 16.3     | 20.0     | 19.8     | 14.7     | 19.6     | 22.5     | 24.7     | 19.225    | 25.326 |          |           |         |      |      |      |      |
| 7   | +1                    | -1       | +1       | 0        | 16.4                  | 19.1     | 18.4     | 23.6     | 16.8     | 18.6     | 24.3     | 21.6     | 19.850    | 25.711 |          |           |         |      |      |      |      |
| 8   | +1                    | 0        | -1       | +1       | 14.2                  | 15.6     | 15.1     | 16.8     | 17.8     | 19.6     | 23.2     | 24.2     | 18.338    | 24.852 |          |           |         |      |      |      |      |
| 9   | +1                    | +1       | 0        | -1       | 16.1                  | 19.9     | 19.3     | 17.3     | 23.1     | 22.7     | 22.6     | 28.6     | 21.200    | 26.152 |          |           |         |      |      |      |      |

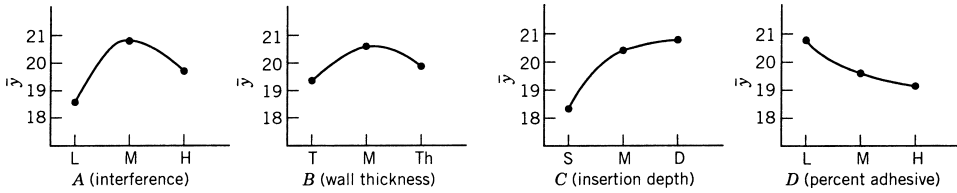


Figure 11.5 Plot of  $\bar{y}$  against levels of control variables.

As in Example 11.1, the graphs are examined and the analysis involves a pick-the-winner approach. The following represent the optimal choices of  $A$ ,  $B$ ,  $C$ , and  $D$  according to the plots:

- $A$  : Medium
- $B$  : Medium
- $C$  : Medium or deep
- $D$  : Low

In each case, we have the levels that maximize the mean response and also maximize  $SNR_I$ . Again, these choices are made under the assumption that no interaction exists among the control factors.

Actually, when cost and other external factors were taken into account, the researchers decided to use

- $A$  : Medium
  - $B$  : Thin
  - $C$  : Medium
  - $D$  : Low
- (11.6)

The thin level for  $B$  was much less expensive. Neither of the two conditions for  $A$ ,  $B$ ,  $C$ , and  $D$  reported above was in the design. As a result, confirmatory runs were necessary. The authors report that good results were observed. The levels used in the noise variables were  $E_{low}$ ,  $F_{low}$ , and  $G_{low}$ .

### 11.3.3 Further Comments

To this point our emphasis has been on experimental design and analysis procedures suggested by Taguchi. It should be emphasized that despite criticisms put forth by statisticians

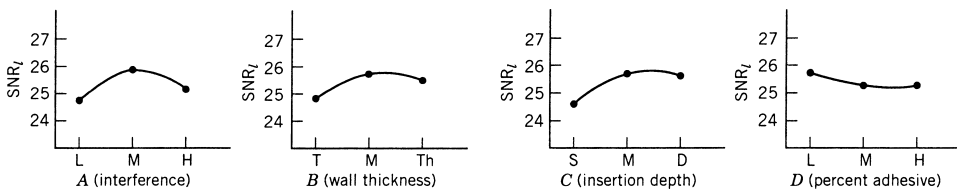


Figure 11.6 Plot of mean  $SNR_I$  against levels of control variables.



concerning his approach, the methods advocated by Taguchi have been successful, and many practitioners have used them. While there are weaknesses in the methodology as illustrated in the two previous examples, there are certainly many real-life situations in which it works well. For an excellent account of Taguchi's contributions the reader is referred to Pignatiello and Ramberg (1991). In the sections that follow, we will discuss areas where there are weaknesses and will point out situations in which these weaknesses may lead to difficulties. This is designed to motivate an emphasis on response surface methods as an alternative for solving robust parameter design problems.

**Experimental Design** The **crossed array**, or **product array**, begins with two experimental designs, one for the noise variables and one for the control variables. These designs are then crossed with each other. The two individual designs are generally quite economical because they are either saturated or near-saturated. However, the crossing of two designs often does not produce an economical design. Indeed, often the designs are huge, as in Example 11.2, where 72 runs were used to study only seven factors. As we will suggest in a subsequent section, there are more economical approaches to designing an experiment for control and noise variables.

In sections that follow we suggest economical experimental designs for handling robust design problems in which control  $\times$  control interactions are available. In addition, these designs are compatible with a response surface approach to the solution of robust parameter design problems.

**The Taguchi Analysis** Notwithstanding difficulties discussed earlier with the SNRs, the approach taken by Taguchi and described here can work very well for determining a robust product if control  $\times$  control interactions are not important. However, the marginal means approach, while placing emphasis on process optimization, does not allow the practitioner to learn more about the process. Lack of an empirical model is a major flaw in the approach. Here, the focus is on models that involve control  $\times$  control and control  $\times$  noise interactions and on the determining of separate models for the process mean and variance.

## 11.4 THE RESPONSE SURFACE APPROACH

There are several response surface alternatives for solving the RPD problem and for conducting process robustness studies. In this section we begin by emphasizing the importance of the control  $\times$  noise interactions, discuss alternative modeling and analysis strategies, and present alternatives to the Taguchi-style crossed array experimental designs. With regard to modeling strategies, it should be emphasized that, in general, process optimization may be overrated. The more emphasis is placed on learning about the process, the less important *absolute optimization* becomes. An engineer or scientist always can (and often does) find pragmatic reasons why the suggested optimum conditions cannot be adopted. A thorough study (via a response surface approach) of the control variable design space will generally provide effective alternative suggestions for operating conditions.

### 11.4.1 The Role of the Control $\times$ Noise Interaction

We have emphasized the need to model control  $\times$  control interactions. However, the control  $\times$  noise interactions are vitally important. Indeed, the structure of these interactions determines the nature of nonhomogeneity of process variance that characterizes the

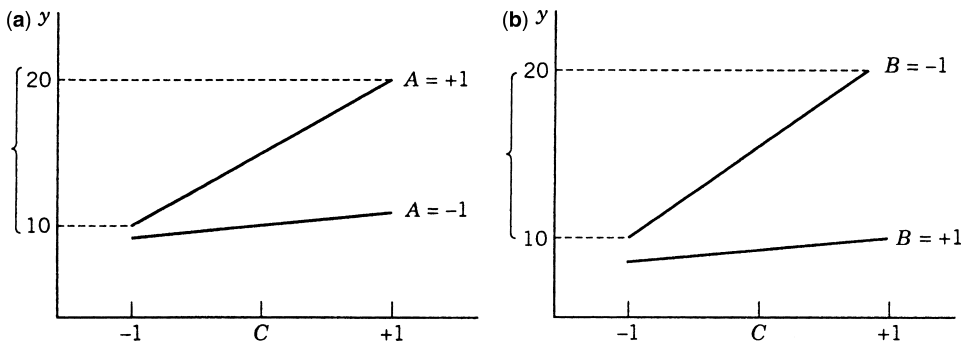
**TABLE 11.5 Experimental Data in a Crossed Array**

| Inner Array |    | Outer Array |    | Response Means |
|-------------|----|-------------|----|----------------|
| A           | B  | C = -1      | +1 |                |
| -1          | -1 | 11          | 15 | 13.0           |
| -1          | 1  | 7           | 8  | 7.5            |
| 1           | -1 | 10          | 26 | 18.0           |
| 1           | 1  | 10          | 14 | 12.0           |

parameter design problem. As an illustration consider the case of two control variables and one noise variable with data as shown in the crossed array design of Table 11.5.

Consider the plots displaying the control  $\times$  noise interactions  $AC$  and  $BC$  in Figs. 11.7a and 11.7b. In viewing these interaction plots, keep in mind that while  $C$  is held fixed in the experimental design, it will vary according to some probability distribution during the process. The probability distribution may be discrete or continuous. Here it is clear that when we take this variability into account, the picture suggests that  $A = +1$  and  $B = -1$  will result in *largest product variability*, whereas  $A = -1$  and  $B = +1$  should result in *smallest product variability*. The braces on the vertical axes reflect product variability at  $A = +1$  and  $B = -1$ . Now, of course, in this simple case, the conclusion is borne out in the data themselves, but it may not be so easy to see in larger problems. In this illustration, we say that  $A$  and  $B$  are both **dispersion effects**, that is, effects that have influence on the process variance. In addition, both variables have **location effects**, namely, they have an effect on the process mean. In fact, the illustration points out the need to compromise between process mean and process variance. From the plots, the combination  $A = +1$  and  $B = -1$  produces the largest mean response, and if this is a larger-the-better case, determining the optimum conditions will involve some type of simultaneous optimization of the process mean and variance.

It should be apparent that control  $\times$  noise interactions display the structure of the process variance and relate which factors are dispersion effects and which are not. Consider the example in Table 11.6. The  $A \times C$  and  $B \times C$  interaction plots are given in Figs. 11.8a and 11.8b. Here we have parallelism in both interaction plots. There is no  $AC$  factor interaction



**Figure 11.7** Interaction plots for the data in Table 11.5. (a)  $AC$  interaction plot. (b)  $BC$  interaction plot.

**TABLE 11.6** Experimental Data in a Crossed Array

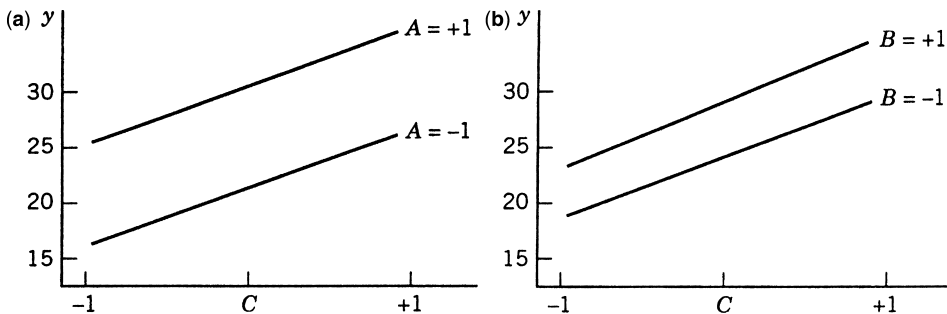
| Inner Array |     | Outer Array |      |
|-------------|-----|-------------|------|
| $A$         | $B$ | $C = C-1$   | $+1$ |
| -1          | -1  | 14.5        | 22.5 |
| -1          | 1   | 19          | 27.5 |
| 1           | -1  | 23.5        | 32   |
| 1           | 1   | 28.5        | 37   |

and consequently  $A$  is *not* a dispersion effect. That is, the process variance is the same at both levels of  $A$ . There is no  $BC$  interaction, so  $B$  is *not* a dispersion effect either. The factor  $A$  has a greater location effect than  $B$ . If this is a larger-the-better case, then the combination  $A = +1$  and  $B = +1$  is optimal and there is *no robust parameter design problem*, because altering  $A$  and  $B$  has no effect on the variance produced by changing the noise variable  $C$ .

Consider a third example, for which the  $AC$  and  $BC$  interaction plots appear in Figs. 11.9a and 11.9b. Figure 11.9a depicts  $A$  as a dispersion effect with  $A = +1$  representing the level of minimum variance. Figure 11.9b suggests that  $B$  is *not* a dispersion effect. Both  $A$  and  $B$  are location effects; and for a larger-the-better scenario,  $A = +1$  and  $B = +1$  represents optimum conditions. The control  $\times$  noise interactions can serve as nice diagnostic tools regarding the structure of the process variance. Note that the larger the slopes of the interaction plots, the larger the contributions to the process variance. Here we have discussed the diagnostic nature of the two-factor interactions. However, if there are three-factor interactions involving control and noise factors, these too play a role.

Consider the connector pulloff force problem in Example 11.2. Here, there are four control factors; from Fig. 11.6 factor  $A$  appears to have an effect on  $SNR_j$ , and indeed we learned that  $A$  medium is the appropriate level. In addition,  $A$  medium maximizes  $\bar{y}$ . The two-factor  $AG$  interaction plot is given in Fig. 11.10. This plot implies that  $A$  medium *minimizes the variance produced by changes in relative humidity*. It also reinforces that  $A$  medium is desirable from the point of view of the mean response.

As a result, it should be clear that the control  $\times$  noise interaction plot aids in the determination of which factors are dispersion effects, which are location effects, and which



**Figure 11.8** Interaction plots for the data in Table 11.6. (a)  $AC$  interaction plot. (b)  $BC$  interaction plot.

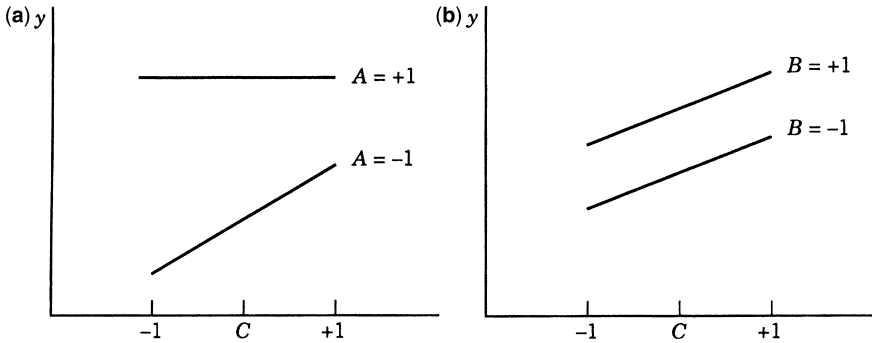


Figure 11.9 Two-factor interaction plot. (a) AC interaction plot. (b) BC interaction plot.

levels of the dispersion effects moderate the variance produced by which noise variables. For example, consider a situation in which we have three control factors ( $A$ ,  $B$ , and  $C$ ) and two noise factors ( $D$  and  $E$ ). Suppose also that interactions  $AD$  and  $BE$  are found to be significant. In addition,  $A$  is found to be insignificant while  $B$  and  $C$  are significant.

Consider the main effect and interaction plots in Fig. 11.11. The plots indicate that  $C$  has a strong location effect and  $B$  has a mild location effect. Factor  $B$  has a huge dispersion effect with minimum process variance resulting from the use of the  $-1$  level of  $B$ . Using the  $-1$  level of  $B$  minimizes that portion of the process variance that results from changes in the noise variable  $E$ . Despite the fact that  $A$  interacts with the noise variable  $D$ , the use of  $-1$  or  $+1$  on  $A$  results in the same process variance. However, if  $A$  is a continuous variable, there may be other levels of  $A$  that will produce a horizontal line on the  $AD$  interaction plot and thus dramatically reduce the portion of the process variance associated with changes in  $D$ . (Note the dashed line on the  $AD$  interaction plot.) Because the factor  $C$  does not interact with either noise factor,  $C$  cannot be used to control process variance. If one were

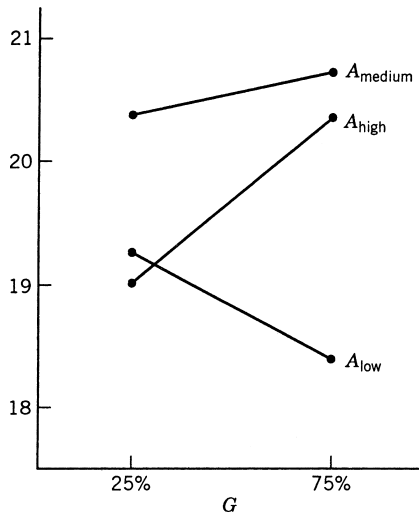


Figure 11.10 AG interaction plot for Example 11.2.

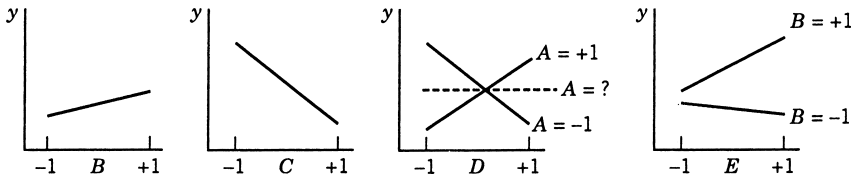


Figure 11.11 Main effect and interaction plots.

interested in conditions that *minimize mean response*, say, but also result in minimizing process variance, one would choose  $B = -1$ ,  $C = +1$ , and  $A = -1$  or  $+1$ . However, if  $A$  is a continuous variable, the value of  $A$  that results in minimum process variance should be sought. This latter idea will be developed in future sections.

**11.4.2 A Model Containing Both Control and Noise Variables**

From the discussion in the previous subsection, it becomes apparent that one may wish to make use of a model that contains main effects and interactions involving both control and noise variables. If the control variables are predominantly continuous variables, this type of model will become a very useful vehicle for the eventual formation of a **dual response surface** approach—that is, a response surface for the process mean and a response surface for the process variance. In what has preceded this section, we have assumed that control and noise factors are fixed effects. Suppose we consider  $\mathbf{x}$  and  $\mathbf{z}$ , the control and noise factors respectively. The modeling of both  $\mathbf{x}$  and  $\mathbf{z}$  in the same model has been called a **response model approach**. The following example illustrates one approach to the analysis of the response model.

**Example 11.3 The Pilot Plant Experiment** Consider the experiment described in Example 3.2, where a  $2^4$  factorial design was used to study the filtration rate of a chemical product. The four factors are temperature, pressure, concentration of formaldehyde, and stirring rate. Each factor is present at two levels. The design matrix and the response data obtained from a single replicate of the  $2^4$  experiment are repeated for convenience in Table 11.7. The 16 runs were made in random order.

Assume that temperature is hard to control in production. Therefore, it is the noise variable and is denoted by  $z_1$ . The other three factors are denoted  $x_1 =$  pressure,  $x_2 =$  concentration, and  $x_3 =$  stirring rate. Because both the control factors and the noise factor are in the same design matrix, this type of design is called a **combined array design**.

Using the results from the original analysis in Example 3.2, the response model is

$$\begin{aligned} \hat{y}(\mathbf{x}, z_1) &= 70.06 + \left(\frac{21.625}{2}\right)z_1 + \left(\frac{9.875}{2}\right)x_2 + \left(\frac{14.625}{2}\right)x_3 \\ &\quad - \left(\frac{18.125}{2}\right)x_2z_1 + \left(\frac{16.625}{2}\right)x_3z_1 \\ &= 70.06 + 10.81z_1 + 4.94x_2 + 7.31x_3 - 9.06x_2z_1 + 8.31x_3z_1 \end{aligned}$$

TABLE 11.7 Pilot Plant Filtration Rate Experiment

| Run Number | Factor |       |       |       | Filtration Rate (gal/hr) |
|------------|--------|-------|-------|-------|--------------------------|
|            | $z_1$  | $x_1$ | $x_2$ | $x_3$ |                          |
| 1          | -      | -     | -     | -     | 45                       |
| 2          | +      | -     | -     | -     | 71                       |
| 3          | -      | +     | -     | -     | 48                       |
| 4          | +      | +     | -     | -     | 65                       |
| 5          | -      | -     | +     | -     | 68                       |
| 6          | +      | -     | +     | -     | 60                       |
| 7          | -      | +     | +     | -     | 80                       |
| 8          | +      | +     | +     | -     | 65                       |
| 9          | -      | -     | -     | +     | 43                       |
| 10         | +      | -     | -     | +     | 100                      |
| 11         | -      | +     | -     | +     | 45                       |
| 12         | +      | +     | -     | +     | 104                      |
| 13         | -      | -     | +     | +     | 75                       |
| 14         | +      | -     | +     | +     | 86                       |
| 15         | -      | +     | +     | +     | 70                       |
| 16         | +      | +     | +     | +     | 96                       |

Notice that the fitted response model contains both the control and the noise factors. Let us assume that  $z_1$  is a random variable with mean zero and variance  $\sigma_z^2$ . Also, assume that  $\sigma_z^2$  is known and the high and low levels of  $z_1$  are at  $\pm\sigma_z$  in coded form. Thus,  $\sigma_z = 1$ .

Suppose that the fitted model above is adequate to allow us to assume that the true relationship between  $y$ ,  $\mathbf{x}$ , and  $z_1$  is

$$y(\mathbf{x}, z_1) = \beta_0 + \gamma_1 z_1 + \beta_2 x_2 + \beta_3 x_3 + \delta_{12} z_1 x_2 + \delta_{13} z_1 x_3 + \varepsilon$$

Because  $E(z_1) = 0$  and  $E(\varepsilon) = 0$ , a model for the **process mean** is

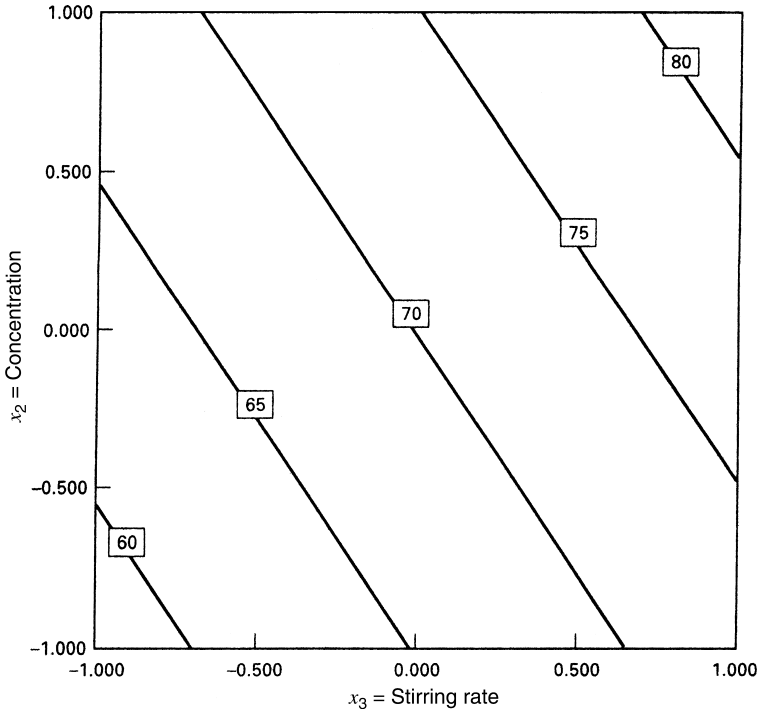
$$E_{z_1}[y(\mathbf{x}, z_1)] = \beta_0 + \beta_2 x_2 + \beta_3 x_3$$

Now apply a conditional variance operator across  $y(\mathbf{x}, z_1)$ . This gives

$$\begin{aligned} \text{Var}_{z_1}[y(\mathbf{x}, z_1)] &= \text{Var}_{z_1}[\beta_0 + \gamma_1 z_1 + \beta_2 x_2 + \beta_3 x_3 + \delta_{12} z_1 x_2 + \delta_{13} z_1 x_3 + \varepsilon] \\ &= \text{Var}_{z_1}[\beta_0 + \beta_2 x_2 + \beta_3 x_3 + (\gamma_1 + \delta_{12} x_2 + \delta_{13} x_3) z_1 + \varepsilon] \\ &= \sigma_z^2 (\gamma_1 + \delta_{12} x_2 + \delta_{13} x_3)^2 + \sigma^2 \end{aligned}$$

This is a model for the **process variance**. We can replace the parameters by their estimates to obtain

$$\widehat{E_z}[y(\mathbf{x}, z_1)] = 70.06 + 4.94x_2 + 7.31x_3$$



**Figure 11.12** Contours of constant mean filtration rate (Example 11.3) with  $z_1 = \text{temperature} = 0$ .

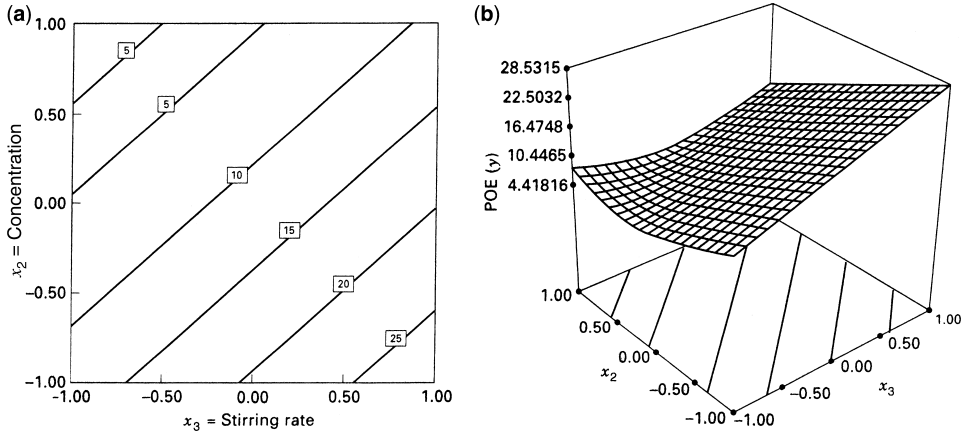
and

$$\begin{aligned} \widehat{\text{Var}}_z[y(\mathbf{x}, z_1)] &= \sigma_z^2(10.81 - 9.06x_2 + 8.31x_3)^2 + \sigma^2 \\ &= 136.42 + 82.08x_2^2 + 69.06x_3^2 - 195.88x_2 + 179.66x_3 - 150.58x_2x_3 \end{aligned}$$

because  $\sigma_z^2=1$  and  $\hat{\sigma}^2 = 19.51$  (this is the residual mean square obtained by fitting the response model).

Figure 11.12 presents a contour plot of the response contours from the mean model. This plot was constructed using Design-Expert. To construct this plot, we held the noise factor (temperature) at zero and the nonsignificant controllable factor (pressure) at zero. Notice that the mean filtration rate increases as either the concentration or the stirring rate increases. Design-Expert will also construct plots automatically of the *square root* of the variance contours, which it labels **propagation of error**, or **POE**. Obviously, the POE is just the standard deviation of the transmitted variability in the response as a function of the controllable variables. Figure 11.13 shows a contour plot and a three-dimensional response surface plot of the POE, obtained from Design-Expert (in this plot the noise variable is held constant at zero, as explained previously).

Suppose that the experimenter wants to maintain a mean filtration rate about 75 and minimize the variability around this value. Figure 11.14 shows an overlay plot of the contours of mean filtration rate and the POE as a function of concentration and stirring rate, the significant controllable variables. To achieve the desired objectives, it will be necessary to hold concentration at the high level and stirring rate very near the middle level.

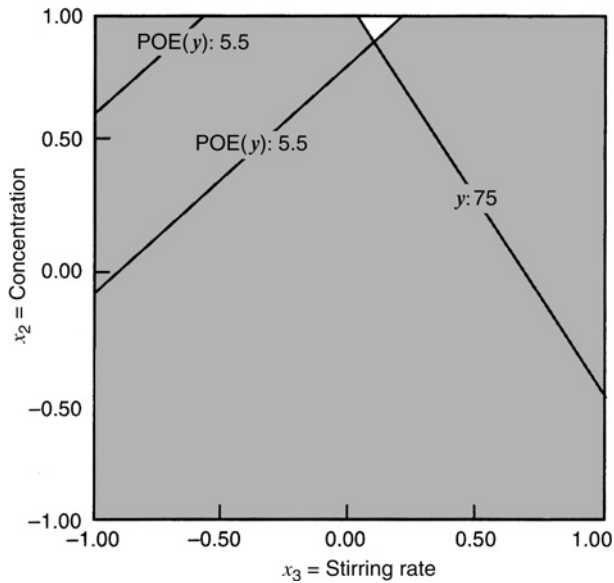


**Figure 11.13** Contour plot and response surface plot of propagation of error for Example 11.3, with  $z_1 = \text{temperature} = 0$ . (a) Contour plot. (b) Response surface plot.

**11.4.3 Generalization of Mean and Variance Modeling**

The analysis in the previous subsection indicates that a *dual* response approach to the RPD problem or a process robustness study offers considerable flexibility to the analyst. Specifically:

1. It provides an estimate of the mean and standard deviation at any location of interest in the control design variables.



**Figure 11.14** Overlay plot of mean and POE contours for filtration rate (Example 11.3), with  $z_1 = \text{temperature} = 0$ .



2. The engineer or scientist can gain insight regarding the roles of these variables in controlling the process mean and variance.
3. It provides a ready source of process optimization via the use of a squared error loss criterion such as

$$E_z[y(\mathbf{x}, \mathbf{z}) - t]^2 = \{E[y(\mathbf{x}, \mathbf{z})] - t\}^2 + \text{Var}_z[y(\mathbf{x}, \mathbf{z})]$$

which will result in a single-number criterion appropriate for the target-is-best scenario.

4. It allows the use of a wide variety of constrained optimization techniques.

Many authors have pointed out the advantages of mean and variance modeling and illustrated its use in both RPD problems and process robustness studies. Some useful references include, Box and Jones (1989, 1992), Vining and Myers (1990), Shoemaker, Tsui, and Wu (1991), Myers, Khuri, and Vining (1992a), Del Castillo and Montgomery (1993), Lucas (1994), Pledger (1996), Myers, Kim, and Griffiths (1997), Khattree (1996), Montgomery (1999), and Borror and Montgomery (2000).

The example in Section 11.4.2 involved a first-order model in the control factors. In many situations, the model in the control and noise variables may involve **quadratic**, or second-order, terms in the control variables (the  $x$ 's). Thus a plausible model would be

$$y(\mathbf{x}, \mathbf{z}) = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x} + \mathbf{z}'\boldsymbol{\gamma} + \mathbf{x}'\boldsymbol{\Delta}\mathbf{z} + \varepsilon \quad (11.7)$$

In this **response model** we assume linear effects in  $\mathbf{x}$  and  $\mathbf{z}$ , two-factor interaction and pure quadratic terms in  $\mathbf{x}$  (the term  $\mathbf{x}'\mathbf{B}\mathbf{x}$ ), and the very important two-factor interactions involving control and noise variables (the term  $\mathbf{x}'\boldsymbol{\Delta}\mathbf{z}$ ).

We do not rule out the use of interactions among noise variables or even quadratic terms in noise variables. However, the model in Equation (11.7) accommodates many real-life situations. The development of mean and variance modeling with the use of Equation 11.7 provides the machinery for more general applications. Let  $\mathbf{z}$  in Equation 11.7 be  $r_z \times 1$ , and let  $\mathbf{x}$  be  $r_x \times 1$ . Then  $\boldsymbol{\Delta}$  is an  $r_x \times r_z$  matrix containing coefficients of the control  $\times$  noise or  $x \times z$  interactions.

**Assumptions Underlying Dual Response Development** The  $\varepsilon$ -term in Equation 11.7 plays the same role as the random error term in any response surface methodology (RSM) model discussed in Chapters 3–10. Assumptions that  $\varepsilon_i$  are NID(0,  $\sigma^2$ ) are often reasonable. It should be pointed out that the *constant variance* assumption on  $\varepsilon$  implies that any nonconstancy of variance of the process stems primarily from *inability to control noise variables*. Detection of other sources of heterogeneous variance will be addressed in Section 11.6.

The assumptions associated with the noise variables deserve special attention. As we indicated in Example 11.3, we must assume that the user possesses a reasonable amount of knowledge about the random variable  $\mathbf{z}$ . It may be assumed that in the experiment the natural levels of  $z_j$  are centered at the mean that  $z_j$  experiences in the process and that the  $\pm 1$  levels are at  $\mu_{z_j} + \sigma_{z_j}$ . As a result, the mean of the  $z$ 's in coded metric is at 0, and  $\pm 1$  levels are one standard deviation from the mean. The latter is not a hard and fast assumption. The experimenter may decide to let  $\pm 1$  correspond to  $\pm 2\sigma_z$ . Thus, we will demonstrate

a development of the process variance response surface under the assumption that  $\sigma_z$  is known, and the latter may take on values of  $1, \frac{1}{2}, \frac{1}{3}$  and so on, in practice. For this purpose, assume that

$$E(\mathbf{z}) = \mathbf{0}, \quad \text{Var}(\mathbf{z}) = \sigma_z^2 \mathbf{I}_{r_z} \quad (11.8)$$

Equation 11.8 implies the assumption that noise variables are uncorrelated with known variance.

**The Mean and Variance Response Surfaces** We now apply conditional expectation and variance operators to the model in Equation 11.8. Because  $E(\mathbf{z}) = \mathbf{0}$ , the expectation operation  $E_z$  gives the response surface model for the mean:

$$E_z[y(\mathbf{x}, \mathbf{z})] = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x} \quad (11.9)$$

The conditional variance operator,  $\text{Var}_z[y(\mathbf{x}, \mathbf{z})]$  plus the error variance around Equation 11.7 gives

$$\text{Var}_z[y(\mathbf{x}, \mathbf{z})] = \text{Var}_z(\boldsymbol{\gamma}' + \mathbf{x}'\boldsymbol{\Delta})\mathbf{z} + \sigma^2$$

Here the quantity  $\boldsymbol{\gamma}' + \mathbf{x}'\boldsymbol{\Delta} = \mathbf{a}'$  (say) is a vector of constants. As a result, we have to evaluate  $\text{Var}_z(\mathbf{a}'\mathbf{z})$ . Rules for applying the variance operator give

$$\text{Var}_z(\mathbf{a}'\mathbf{z}) = \mathbf{a}'\text{Var}(\mathbf{z})\mathbf{a}$$

where  $\text{Var}(\mathbf{z}) = \sigma_z^2 \mathbf{I}$  is the variance-covariance matrix of  $\mathbf{z}$ . As a result, we have

$$\text{Var}_z(\boldsymbol{\gamma}' + \mathbf{x}'\boldsymbol{\Delta})\mathbf{z} = \sigma_z^2(\boldsymbol{\gamma}' + \mathbf{x}'\boldsymbol{\Delta})(\boldsymbol{\gamma}' + \mathbf{x}'\boldsymbol{\Delta})'$$

so

$$\begin{aligned} \text{Var}_z[y(\mathbf{x}, \mathbf{z})] &= \sigma_z^2(\boldsymbol{\gamma}' + \mathbf{x}'\boldsymbol{\Delta})(\boldsymbol{\gamma}' + \mathbf{x}'\boldsymbol{\Delta})' + \sigma^2 \\ &= \sigma_z^2 \mathbf{l}'(\mathbf{x})\mathbf{l}(\mathbf{x}) + \sigma^2 \end{aligned} \quad (11.10)$$

where  $\mathbf{l}(\mathbf{x}) = \boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{x}$ .

Equation 11.9 is the **response surface for the process mean**, while Equation 11.10 is **the response surface for the process variance**. Equation 11.10 deserves special attention. Apart from  $\sigma^2$ , the model error variance, the process variance is  $\|\mathbf{l}(\mathbf{x})\|^2$ , where  $\mathbf{l}(\mathbf{x})$  is simply the vector of partial derivatives of  $y(\mathbf{x}, \mathbf{z})$  with respect to  $\mathbf{z}$ . In other words,

$$\mathbf{l}(\mathbf{x}) = \frac{\partial y(\mathbf{x}, \mathbf{z})}{\partial \mathbf{z}}$$

is just the **slope** of the response surface in the direction of the noise variables. This is very intuitive: the larger the vector of derivatives  $\partial y/\partial \mathbf{z}$  (the slopes), the larger the process variance. This relates very nicely to the discussion in Section 11.4.1. Recall that the flatter the  $x \times z$  interaction plots, the smaller the process variance. Clearly, *flat interaction plots result from small values of  $\partial y(\mathbf{x}, \mathbf{z})/\partial \mathbf{z}$* . Also, note the important role of  $\boldsymbol{\Delta}$ , the matrix of  $x \times z$

interaction coefficients, in the process variance response surface. If  $\Delta = \mathbf{0}$ , process variance does not depend on  $\mathbf{x}$ , and hence one cannot create a robust process by choice of settings of the control variables. In other words, the process variance is constant and thus *there is no robust parameter design problem*. In Equation 11.10 we add the model error variance. In many cases this variance will be negligible compared to the variance that is contributed by variation in the noise variables. In our development here there is an implicit assumption that the noise variables are quantitative in nature. In the case of discrete noise variables the assumption of uncorrelated noise variables is not necessarily appropriate.

Equations 11.9 and 11.10 represent the mean and variance response surfaces developed from the model (or response surface) that contains both control and noise variables Equation 11.7. Clearly the estimated response surfaces are obtained by replacing parameters by estimates found in the fitted model and

$$\hat{y}(\mathbf{x}, \mathbf{z}) = b_0 + \mathbf{x}'\hat{\mathbf{b}} + \mathbf{x}'\hat{\mathbf{B}}\mathbf{x} + \mathbf{z}'\hat{\mathbf{c}} + \mathbf{x}'\hat{\mathbf{\Delta}}\mathbf{z} \quad (11.11)$$

Consequently, the estimated process mean and variance response surfaces are given by

$$\begin{aligned} \widehat{E_z[y(\mathbf{x}, \mathbf{z})]} &= \hat{\mu}_z[y(\mathbf{x}, \mathbf{z})] \\ &= b_0 + \mathbf{x}'\hat{\mathbf{b}} + \mathbf{x}'\hat{\mathbf{B}}\mathbf{x} \end{aligned} \quad (11.12)$$

$$\begin{aligned} \widehat{\text{Var}_z[y(\mathbf{x}, \mathbf{z})]} &= \hat{\sigma}_z^2 \hat{\mathbf{I}}'(\mathbf{x})\hat{\mathbf{I}}(\mathbf{x}) + \hat{\sigma}^2 \\ &= \hat{\sigma}_z^2 (\mathbf{c}' + \mathbf{x}'\hat{\mathbf{\Delta}})(\mathbf{c} + \hat{\mathbf{\Delta}}'\mathbf{x}) + \hat{\sigma}^2 \end{aligned} \quad (11.13)$$

Here,  $\hat{\sigma}^2$  is the mean squared error in the fitted response surface.

It is actually possible to generalize these results a bit more. We may write the response model involving the control and noise variables as

$$y(\mathbf{x}, \mathbf{z}) = f(\mathbf{x}) + h(\mathbf{x}, \mathbf{z}) + \varepsilon \quad (11.14)$$

where  $f(\mathbf{x})$  is the portion of the model that involves only the control variables and  $h(\mathbf{x}, \mathbf{z})$  are the terms involving the noise factors and the interactions between the controllable and noise factors. In Equation 11.7, the structure for  $h(\mathbf{x}, \mathbf{z})$  was assumed to be

$$h(\mathbf{x}, \mathbf{z}) = \sum_{i=1}^{r_z} \gamma_i z_i + \sum_{i=1}^{r_x} \sum_{j=1}^{r_z} \delta_{ij} x_i z_j \quad (11.15)$$

The structure for  $f(\mathbf{x})$  and  $h(\mathbf{x}, \mathbf{z})$  will depend on what type of model the experimenter thinks is appropriate. The logical choices for  $f(\mathbf{x})$  are the first-order model with interaction and the second-order model. Equation 11.15 represents a reasonable choice for  $h(\mathbf{x}, \mathbf{z})$ ; if appropriate, however, one could include more complex functions of the noise variables. If we assume that the noise variables have mean zero, variance  $\sigma_z^2$ , and covariance zero, and that the noise variables and the random errors  $\varepsilon$  have zero covariances, then the mean model for the response is just

$$E_z[y(\mathbf{x}, \mathbf{z})] = f(\mathbf{x}) \quad (11.16)$$

and the variance model for the response is

$$\text{Var}_z[y(\mathbf{x}, \mathbf{z})] = \sigma_z^2 \sum_{i=1}^r \left[ \frac{\partial y(\mathbf{x}, \mathbf{z})}{\partial z_i} \right]^2 + \sigma^2 \quad (11.17)$$

Equation 11.17 was found by expanding  $y(\mathbf{x}, \mathbf{z})$  in a first-order Taylor series about  $\mathbf{z} = 0$  and applying the variance operator (this is sometimes called the **delta method**). Equation 11.17 is also called the **transmission-of-error formula**. When the noise variables are involved linearly in  $h(\mathbf{x}, \mathbf{z})$ , Equation 11.17 is exactly the same as Equation 11.10.

#### 11.4.4 Analysis Procedures Associated with the Two Response Surfaces

We saw a simple special case of a dual response surface analysis in Section 11.4.2. We were able to find operating conditions by overlaying the response surfaces. Another approach is to use constrained optimization, that is, select a value of  $E_z[y(\mathbf{x}, \mathbf{z})]$  that one is willing to accept. This does not imply a target on the mean but, say, a value below which one cannot accept. In fact, several of these values may be chosen in order to add flexibility. Call this constraint  $E_z[y(\mathbf{x}, \mathbf{z})] = m$ . One then chooses  $\mathbf{x}$  as follows:

$$\text{Min}_{\mathbf{x}} \text{Var}_z[y(\mathbf{x}, \mathbf{z})] \quad \text{subject to} \quad E_z[y(\mathbf{x}, \mathbf{z})] = m$$

Several values of  $m$  may be used in order to provide flexibility for the user.

There are many situations in which focus is placed entirely on the **variance** response surface. For example, there may be location effects that appear in the mean model but not in the variance model, and they can be used as tuning factors. As a result, the analysis centers around the variance model. One may be interested in determining  $\mathbf{x}$  that minimizes, or makes zero, the process variance in Equation 11.10. To that end, consider the computation of the locus of points  $\mathbf{x}_0$  for which

$$\mathbf{l}(\mathbf{x}_0) = \mathbf{0} \quad (11.18)$$

or, in terms of the fitted surface,

$$\mathbf{c} + \hat{\mathbf{\Delta}}' \mathbf{x}_0 = \mathbf{0}$$

The condition Equation 11.18 represents  $r_z$  equations in  $r_x$  unknowns. If  $r_x > r_z$ , this result, which produces a zero estimated process variance (apart from the model error variance  $\sigma^2$ ), will result in a line or a plane. If  $r_x = r_z$  there will a single point  $\mathbf{x}_0$  that satisfies Equation 11.18. If  $r_x < r_z$ , Equation 11.18 may not have a solution. Of course a great deal of interest centers on whether or not the locus of points  $\mathbf{x}_0$  passes through the experimental design region in the control variables.

We now present an example involving a second-order response surface model. This example is taken from Montgomery (2013).

**Example 11.4 A Process Robustness Study in Semiconductor Manufacturing** An experiment was run in a semiconductor manufacturing facility involving two controllable variables. The combined array design used by the experimenters is shown in Table 11.8.

**TABLE 11.8 Combined Array Experiment with Two Controllable Variables and Three Noise Variables, Example 11.4**

| Run Number | $x_1$ | $x_2$ | $z_1$ | $z_2$ | $z_3$ | $y$  |
|------------|-------|-------|-------|-------|-------|------|
| 1          | -1    | -1    | -1    | -1    | 1     | 44.2 |
| 2          | 1     | -1    | -1    | -1    | -1    | 30.0 |
| 3          | -1    | 1     | -1    | -1    | -1    | 30.0 |
| 4          | 1     | 1     | -1    | -1    | 1     | 35.4 |
| 5          | -1    | -1    | 1     | -1    | -1    | 49.8 |
| 6          | 1     | -1    | 1     | -1    | 1     | 36.3 |
| 7          | -1    | 1     | 1     | -1    | 1     | 41.3 |
| 8          | 1     | 1     | 1     | -1    | -1    | 31.4 |
| 9          | -1    | -1    | -1    | 1     | -1    | 43.5 |
| 10         | 1     | -1    | -1    | 1     | 1     | 36.1 |
| 11         | -1    | 1     | -1    | 1     | 1     | 22.7 |
| 12         | 1     | 1     | -1    | 1     | -1    | 16.0 |
| 13         | -1    | -1    | 1     | 1     | 1     | 43.2 |
| 14         | 1     | -1    | 1     | 1     | -1    | 30.3 |
| 15         | -1    | 1     | 1     | 1     | -1    | 30.1 |
| 16         | 1     | 1     | 1     | 1     | 1     | 39.2 |
| 17         | -2    | 0     | 0     | 0     | 0     | 46.1 |
| 18         | -2    | 0     | 0     | 0     | 0     | 36.1 |
| 19         | 0     | -2    | 0     | 0     | 0     | 47.4 |
| 20         | 0     | 2     | 0     | 0     | 0     | 31.5 |
| 21         | 0     | 0     | 0     | 0     | 0     | 30.8 |
| 22         | 0     | 0     | 0     | 0     | 0     | 30.7 |
| 23         | 0     | 0     | 0     | 0     | 0     | 31.0 |

The design is a 23-run variation of a central composition design that was created by starting with a standard CCD for five factors (the cube portion is a  $2^{5-1}$ ) and deleting the axial runs associated with the three noise variables. This design will support a response model that has a second-order model in the controllable variables, the main effects of the three noise variables, and the interactions between the control and noise factors. The fitted response model is

$$\begin{aligned}\hat{y}(\mathbf{x}, \mathbf{z}) = & 30.37 - 2.9x_1 - 4.13x_2 + 2.60x_1^2 + 2.18x_2^2 + 2.87x_1x_2 + 2.73z_1 \\ & - 2.33z_2 + 2.33z_3 + 0.27x_1z_1 + 0.89x_1z_2 + 2.58x_1z_3 + 2.01x_2z_1 \\ & - 1.43x_2z_2 + 1.56x_2z_3\end{aligned}$$

The mean and variance models are

$$\begin{aligned}\widehat{E}_z[y(\mathbf{x}, \mathbf{z})] &= b_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\hat{\mathbf{B}}\mathbf{x} \\ &= 30.37 - 2.92x_1 - 4.13x_1 + 2.60x_1^2 \\ &\quad + 2.18x_2^2 + 2.87x_1x_2\end{aligned}$$

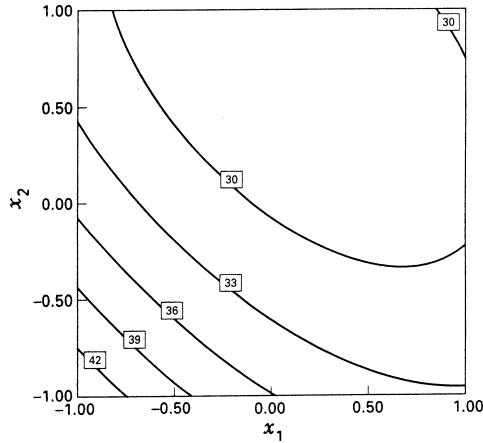


Figure 11.15 Contour plot of the mean model, Example 11.4.

and

$$\begin{aligned} \widehat{\text{Var}}_z[y(\mathbf{x}, \mathbf{z})] &= \hat{\sigma}_z^2 \mathbf{l}'(\mathbf{x})\mathbf{l}(\mathbf{x}) + \sigma^2 \\ &= 19.26 + 3.20x_1 + 12.45x_2 + 7.52x_1^2 + 8.52x_2^2 + 2.21x_1x_2 \end{aligned}$$

where we have substituted parameter estimates from the fitted response model into the equations for the mean and variance models and assumed that  $\sigma_z^2 = 1$ . Figures 11.15 and 11.16 present contour plots of the process mean and POE (remember that the POE is the square root of the variance response surface) generated from these models.

In this problem it is desirable to keep the process mean below 30. From inspection of Figs. 11.15 and 11.16 it is clear that some tradeoff will be necessary if we wish to make the process variance small. Because there are only two controllable variables, a logical way to accomplish this tradeoff is to overlay the contours of constant mean response and constant

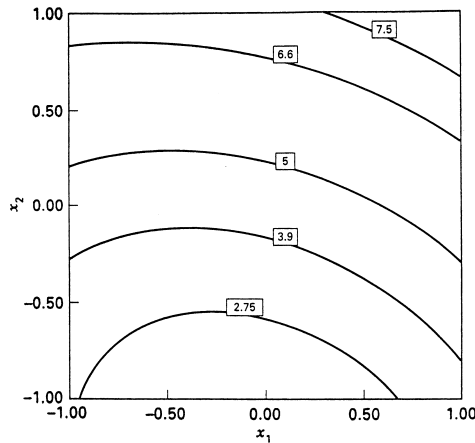
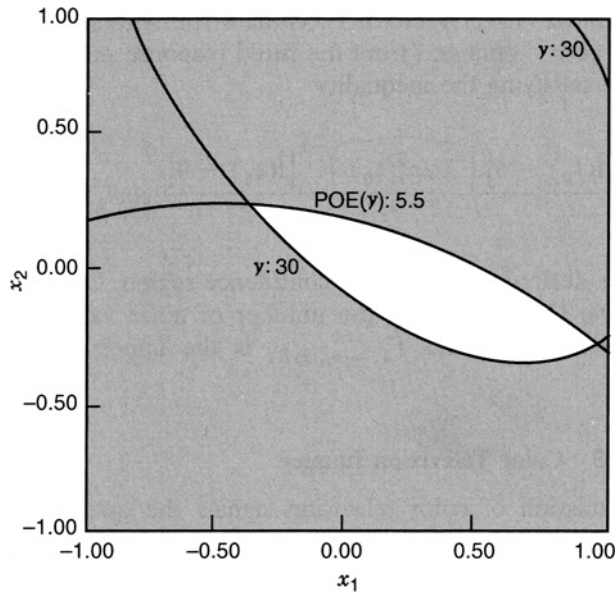


Figure 11.16 Contour plot of the POE, Example 11.4.



**Figure 11.17** Overlay of the mean and POE contours for Example 11.4, with the open region indicating satisfactory operating conditions for process mean and variance.

variance, as shown in Fig. 11.17. This plot shows the contours for which the process mean is less than or equal to 30 and the process standard deviation is less than or equal to 5. The region bounded by these contours would represent a typical operating region of low mean response and low process variance.

There are other approaches to solving the robust design problem in Example 11.4. One approach would be to differentiate the fitted response model with respect to each of the noise variables  $z_1$ ,  $z_2$ , and  $z_3$  (recall that the derivatives of the response model or the slope of the response surface in the direction of the noise variables are proportional to the transmitted variance) and force these derivatives to zero while simultaneously optimizing the levels of the control variables  $x_1$  and  $x_2$ . This approach can be executed in JMP. Figure 11.18a is the output from the JMP prediction profiler showing the response  $y$  and the three derivatives for each of the noise factors after this optimization is carried out using a desirability function technique. JMP also has a simulation capability that predicts the proportion of production outside specifications on the response variable. The proportion of the response  $y$  that is predicted to be above 30 in this problem is about 25%. This indicates that the variability transmitted from the noise variables is still quite high.

In some situations it may be possible to reduce the proportion of unacceptable product by exercising some limited form of control over the noise variables. Often this control takes the form of setting the mean value of the noise variable but still having no control over the variability in the noise variable during routine process operation.

Suppose that in the situation of Example 11.4 we could set the mean of the noise variables, but they would still vary randomly around these means. How much could this additional flexibility potentially reduce the proportion of defective product (where  $y$  is above 30)? The display in Fig. 11.18b addresses this question. This display is the defect profiler from JMP. It provides predictions of how the defect rate is affected by changing

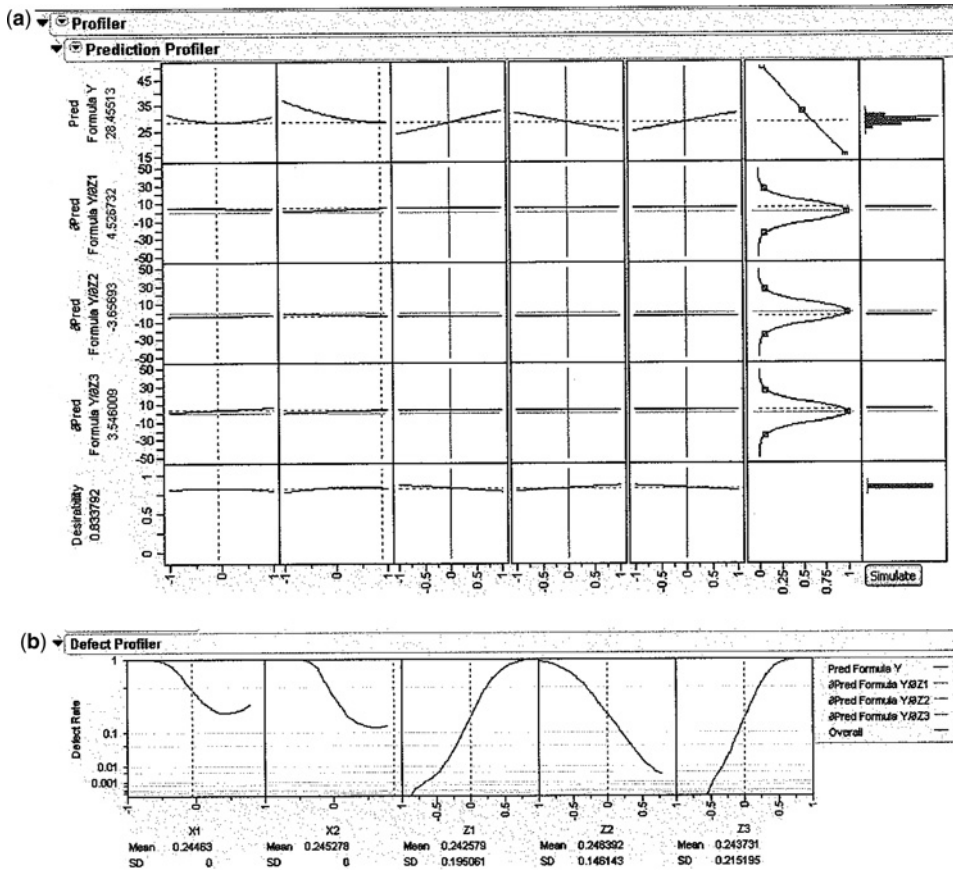


Figure 11.18 Optimal solution to Example 11.4 using the response prediction profiler in JMP. (a) Prediction profiler output from JMP. (b) Defect profiler.

the levels of the design factors. The output indicates that if the mean of the noise variable  $z_1$  can be moved from zero towards its low level, the defect rate will be reduced. Similarly, if the mean of the noise variable  $z_2$  can be moved from zero towards its high level and/or the mean of the noise variable  $z_3$  can be moved from zero towards  $-0.5$ , the defect rate will also be reduced. Therefore, if the experimenter has the additional flexibility of setting the mean or target level of any of the noise variables in the situation, additional improvement in process performance can be obtained.

**Confidence Regions on Minimum Process Variance Conditions** In the previous illustration we displayed a region of low process variance, where the process variance is the variability produced by the lack of sufficient control of noise variables in the process. At times, what is depicted may be a line or a plane or even a point in the space of the control variables. If minimizing process variance is important, then some characterization of the **precision** of the estimates of these conditions may also be crucial. For example, a line of minimum variance may represent a set of conditions that are unsatisfactory to the



investigator. It would be helpful then to gain some sense of how far away from the line the investigator can move and still maintain no significant difference from the minimum variance conditions. The principle here is analogous to that discussed in Chapter 6, where confidence regions were discussed for stationary points in cases where a single response surface is considered. Consider, once again, Equation 11.18 with the solution  $\mathbf{x}_0$  representing a locus of points or a single point of minimum process variance. Let  $\mathbf{t}_0$  represent the true condition of minimum variance—that is,

$$\mathbf{l}(\mathbf{t}_0) = \boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{t}_0 = \mathbf{0}$$

One must keep in mind that  $\mathbf{c}$  is an unbiased estimate of  $\boldsymbol{\gamma}$ , and  $\hat{\boldsymbol{\Delta}}$  is an unbiased estimate of the matrix  $\boldsymbol{\Delta}$ . Now, assuming normal errors around the response model containing  $\mathbf{x}$  and  $\mathbf{z}$ , we have

$$\frac{[\hat{\mathbf{l}}(\mathbf{t}_0) - \mathbf{0}]' [\widehat{\text{Var}} \hat{\mathbf{l}}(\mathbf{t}_0)]^{-1} [\hat{\mathbf{l}}(\mathbf{t}_0) - \mathbf{0}]}{r_z} \sim F_{r_z, df(E)}$$

where  $\hat{\mathbf{l}}(\mathbf{t}_0) = \mathbf{c} + \hat{\boldsymbol{\Delta}}'\mathbf{t}_0$ , and  $\widehat{\text{Var}} \hat{\mathbf{l}}(\mathbf{t}_0)$  is the variance–covariance matrix of  $\hat{\mathbf{l}}(\mathbf{t}_0)$  with  $\sigma^2$  replaced by the error mean square  $s^2$ . See Graybill (1976) or Myers and Milton (1991). Here  $df(E)$  is the error degrees of freedom for the estimator  $s^2$  of  $\sigma^2$  obtained from the fitted response model. As a result the conditions  $\mathbf{t}_0$  satisfying the inequality

$$\frac{[\hat{\mathbf{l}}(\mathbf{t}_0) - \mathbf{0}]' [\widehat{\text{Var}} \hat{\mathbf{l}}(\mathbf{t}_0)]^{-1} [\hat{\mathbf{l}}(\mathbf{t}_0) - \mathbf{0}]}{r_z} \leq F_{1-\alpha, r_z, df(E)} \tag{11.19}$$

represent the desired  $100(1 - \alpha)\%$  confidence region. One should note that  $r_z$  in Equation 11.19 denotes the number of noise variables that interact with control variables, and  $F_{1-\alpha, r_z, df(E)}$  is the upper percentile point of  $F_{r_z, df(E)}$ .

**Example 11.5 Color Television Images** In the transmission of color television signals the quality of the decoded signals is determined by the **PSNRs** (power signal-to-noise ratios in electronics engineering) in the image transmitted. The response here refers to the quality (coarse versus detailed) in the reception of transmitted signals in decibels. The control factors are  $x_1$  (number of tabs in a filter) and  $x_2$  (sampling frequency). These two factors form an inner array, while the noise or outer array involves  $z_1$  (number of bits in an image) and  $z_2$  (voltage applied). The design used is a  $3^2 \times 2^2$  crossed array, and the data appear in Table 11.9.

The model to be used is Equation 11.7. The  $3^2$  in the inner array allows for a second-order model in the control variables. For this specific case, the model can be written as

$$y(\mathbf{x}, \mathbf{z}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \gamma_1 z_1 + \gamma_2 z_2 + \delta_{11} x_1 z_1 + \delta_{12} x_1 z_2 + \delta_{21} x_2 z_1 + \delta_{22} x_2 z_2 + \varepsilon \tag{11.20}$$

**TABLE 11.9 Color TV Image Data for Example 11.5**

|                                    |       |       |       | Codel Levels |         |         |         |
|------------------------------------|-------|-------|-------|--------------|---------|---------|---------|
|                                    |       |       |       | -1           | 0       | 1       |         |
| $x_1$ (number of tabs in a filter) |       |       |       | 5            | 13      | 21      | (tabs)  |
| $x_2$ (sampling frequencies)       |       |       |       | 6.25         | 9.875   | 13.5    | (MHz)   |
| $z_1$ (number of bits of an image) |       |       |       | 256          |         | 512     | (bits)  |
| $z_2$ (voltage applied)            |       |       |       | 100          |         | 200     | (volts) |
| Factor                             |       |       | $z_1$ | -1           | -1      | 1       | 1       |
| Combination                        | $x_1$ | $x_2$ | $z_2$ | -1           | 1       | -1      | 1       |
| (1)                                | -1    | -1    |       | 33.5021      | 41.2268 | 25.2683 | 31.9930 |
| (2)                                | -1    | 0     |       | 35.8234      | 38.0689 | 32.7928 | 34.0383 |
| (3)                                | -1    | 1     |       | 33.0773      | 31.8435 | 36.2500 | 34.0162 |
| (4)                                | 0     | -1    |       | 30.4481      | 41.2870 | 15.1493 | 23.9883 |
| (5)                                | 0     | 0     |       | 34.8679      | 40.2276 | 27.7724 | 31.1321 |
| (6)                                | 0     | 1     |       | 35.2202      | 37.1008 | 33.3280 | 35.2085 |
| (7)                                | 1     | -1    |       | 21.1553      | 34.1086 | 0.7917  | 15.7450 |
| (8)                                | 1     | 0     |       | 27.6736      | 38.1477 | 15.5132 | 25.9873 |
| (9)                                | 1     | 1     |       | 32.1245      | 38.1193 | 26.1673 | 32.1622 |

The process variance is determined by applying the variance operator to Equation 11.20, namely,

$$\begin{aligned}
 \text{Var}_z[y(\mathbf{x}, \mathbf{z})] &= [\gamma_1^2 + \gamma_2^2 + \delta_{11}^2 x_1^2 + \delta_{12}^2 x_1^2 + \delta_{21}^2 x_2^2 + \delta_{22}^2 x_2^2 + 2\gamma_1(\delta_{11}x_1 + \delta_{21}x_2) \\
 &\quad + 2\gamma_2(\delta_{12}x_1 + \delta_{22}x_2) + 2(\delta_{11}\delta_{21} + \delta_{12}\delta_{22})x_1x_2] \sigma_z^2 + \sigma^2 \\
 &= [(\gamma_1 + \delta_{11}x_1 + \delta_{21}x_2)^2 + (\gamma_2 + \delta_{12}x_1 + \delta_{22}x_2)^2] \sigma_z^2 + \sigma^2 \\
 &= \left[ \left( \frac{\partial y(\mathbf{x}, \mathbf{z})}{\partial z_1} \right)^2 + \left( \frac{\partial y(\mathbf{x}, \mathbf{z})}{\partial z_2} \right)^2 \right] \sigma_z^2 + \sigma^2 \\
 &= \{ [l_1(\mathbf{x})]^2 + [l_2(\mathbf{x})]^2 \} \sigma_z^2 + \sigma^2
 \end{aligned}
 \tag{11.21}$$

Note the role of the derivatives  $l_1(\mathbf{x})$  and  $l_2(\mathbf{x})$  in the process variance function. Again, we seek  $\mathbf{x}_0$  that results in minimum process variance. Here we are assuming that the  $\pm$  levels of  $z_1$  and  $z_2$  are at  $\pm\sigma_z$ . Note that the computation of the location of estimated minimum process variance does not depend on knowledge of  $\sigma_z$ .

The least squares fit of the model in Equation 11.20 is given by

$$\begin{aligned}
 \hat{y}(\mathbf{x}, \mathbf{z}) &= 33.389 - 4.175x_1 + 3.748x_2 + 3.348x_1x_2 - 2.328x_1^2 \\
 &\quad - 1.867x_2^2 - 4.076z_1 + 2.985z_2 - 2.324x_1z_1 \\
 &\quad + 1.932x_1z_2 + 3.268x_2z_1 - 2.073x_2z_2
 \end{aligned}$$

All coefficients are significant, and the root error mean square is 0.742 dB. The slopes in the  $z_1$  and  $z_2$  directions are given by

$$\begin{aligned}\hat{l}_1(x_1, x_2) &= -4.076 - 2.234x_1 + 3.268x_2 \\ \hat{l}_2(x_1, x_2) &= 2.985 + 1.932x_1 - 2.073x_2\end{aligned}$$

Both  $x_1$  and  $x_2$  can be used to exert control on the process variance because both control  $\times$  noise interactions are significant.

The point of minimum estimated process variance is found by setting  $\hat{l}_1(\mathbf{x}) = 0$  and  $\hat{l}_2(\mathbf{x}) = 0$ . This gives the result

$$\mathbf{x}_0 = \begin{bmatrix} x_{1,0} \\ x_{2,0} \end{bmatrix} = \begin{bmatrix} -0.874 \\ 0.625 \end{bmatrix}$$

which lies in the experimental design region. For the confidence region around this value, we have

$$\widehat{\text{Var}}[\hat{l}_1(\mathbf{x})] = \widehat{\text{Var}}[\hat{l}_2(\mathbf{x})] = 0.0154 + 0.0230x_1^2 + 0.0230x_2^2$$

This result comes from the variances of the  $\hat{\delta}_{ij}$  and  $c_i$  values. Keep in mind that these estimators are uncorrelated. Thus for the confidence region, Equation 11.20 becomes

$$\frac{\hat{\mathbf{l}}(\mathbf{t})' [\widehat{\text{Var}}(\mathbf{t})]^{-1} \hat{\mathbf{l}}(\mathbf{t})}{2} \leq F_{2,24,0.05}$$

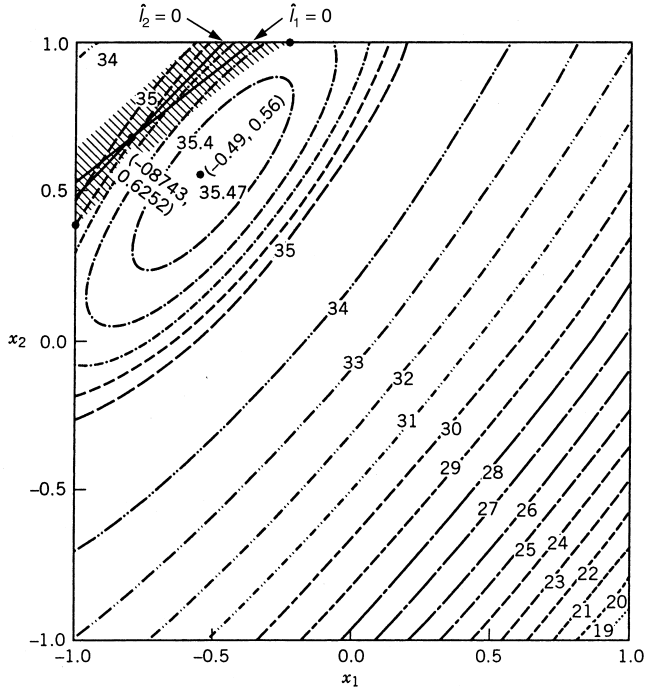
which reduces to

$$\frac{(-4.076 - 2.324t_1 + 3.68t_2)^2 + (2.985 + 1.932t_1 - 2.073t_2)^2}{0.0154 + 0.0230(t_1^2 + t_2^2)} \leq 6.8056$$

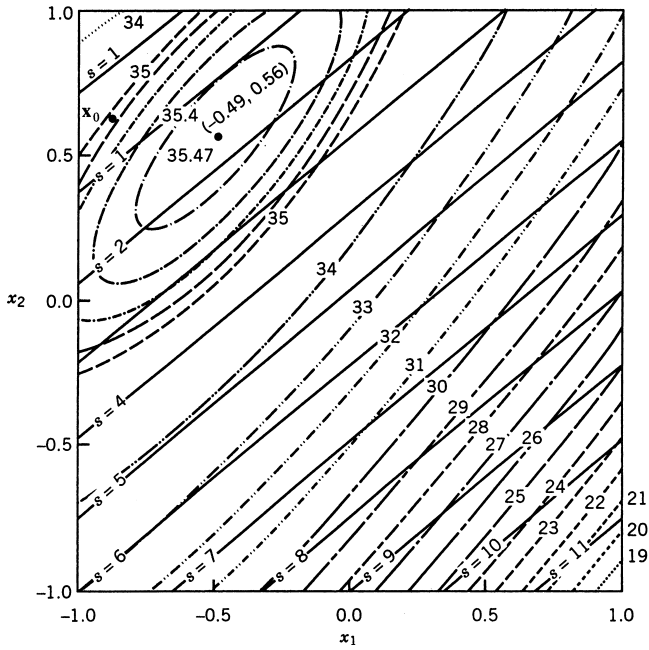
Any  $(t_1, t_2)$  that satisfies the above inequality falls inside the 95% confidence region on the *location of minimum process variance*. See Fig. 11.19. Two lines  $\hat{l}_1(\mathbf{x}) = 0$  and  $\hat{l}_2(\mathbf{x}) = 0$  shows intersection at  $\mathbf{x}_0 = (+0.874, 0.625)$ , the point of estimated minimum process variance. The shaded area displays the confidence region. Also included are the contours of constant *estimated mean response*, whose response surface is obtained using the expectation operator on Equation 11.20:

$$\begin{aligned}\widehat{E}_z[y(\mathbf{x}, \mathbf{z})] &= 33.3889 - 4.1752x_1 + 3.748x_2 \\ &\quad + 3.3485x_1x_2 - 2.3277x_1^2 - 1.8670x_2^2\end{aligned}$$

The notion of a robust product here implies the choice of  $(x_1, x_2)$  that gives *large response with low variance*. Notice that the stationary point for the mean response surface is a point of estimated maximum response. This stationary point, however, does not fall inside the confidence region for the location of minimum variance. The mean response surface is fairly flat in the region of small  $x_1$  and large  $x_2$ . As a result, very little is lost on the mean if one operates at  $\mathbf{x}_0$ , the point of minimum variance. Now, consider Fig. 11.20, where



**Figure 11.19** Contours of constant mean decibels, point of estimated minimum variance, and 95% confidence region on location of process variance for Example 11.5.



**Figure 11.20** Dual response surfaces on mean and standard deviation with points of minimum variance and maximum mean (Example 11.5).

we have superimposed *contours of constant standard deviation*. Here,  $s$  is the estimated process standard deviation. Note that the standard deviation is stable in the upper left-hand corner. The estimated standard deviation at  $\mathbf{x}_0$  is 0.80. Note that the standard deviation becomes less stable as one moves away from the 95% confidence region.

From the display given in Figs. 11.19 and 11.20, the researcher learns a great deal about the process. The most robust process will clearly be in the area of small  $x_1$  and large  $x_2$ . The trough created by the shaded area in Fig. 11.19 shows considerable flexibility as far as optimum conditions are concerned. At  $\mathbf{x}_0$ , the estimated process mean is approximately 35 dB, and the estimated process standard deviation is 0.8.

In this example the experimenter is fortunate that the intersection of  $\hat{l}_1(\mathbf{x}) = 0$  and  $\hat{l}_2(\mathbf{x}) = 0$  occurs inside the design region. If this does not occur, then one must use other response surface methods in dealing with the process variance.

#### 11.4.5 Estimation of the Process Variance

In the previous section we considered the use of a mean and variance response surface for learning about the influence of the control and noise design variables on the process mean and process variance. We also discussed the use of these two response surfaces for finding robust conditions on the control variables, where some type of compromise is chosen for accommodating goals on the mean response but achieving a sufficiently small process variance. It can be argued that the two response surfaces can be dealt with in a fashion similar to that discussed in Chapter 7, where we discussed multiple response surface analysis. It is not our intention to recommend a single “best” way of blending together the mean and variance response surfaces.

In many instances the procedure of determining the conditions that produce zero value for the derivatives (i.e., values for which  $\hat{l}_1 = 0$ ,  $\hat{l}_2 = 0$ , ...,  $\hat{l}_{r_z} = 0$ ) give coordinates that are not practical. There is no assurance that this stationary point will fall inside the region of the design in the control variables. As a result, **ridge analysis** is a reasonable approach for handling both the mean and variance response surfaces. This could provide two loci of points, each locus giving constrained optimum conditions, one locus for the mean and one for the variance. In what follows we will supply a few details for the case of the variance response surface. Before we embark on this discussion we will deal with the notion of appropriate estimation of the process variance.

For the case where the noise variables are uncorrelated, the variance response surface is given by Equation 11.10. In a more general setting [say, when the noise variables are correlated and  $\text{Var}(\mathbf{z}) = \mathbf{V}$ , where  $\mathbf{V}$  is an  $r_z \times r_z$  variance–covariance matrix], the process variance is

$$\text{Var}_z[y(\mathbf{x}, \mathbf{z})] = \mathbf{l}'(\mathbf{x})\mathbf{V}\mathbf{l}(\mathbf{x}) + \sigma^2$$

where  $\sigma^2$  is the error variance for the model of Equation 11.7. Now, if one merely replaces  $\mathbf{l}(\mathbf{x})$  by the estimator from the data, and  $\sigma^2$  by the mean squared error from the fitted model,  $s^2$ , we have as an estimator of the process variance

$$\widehat{\text{Var}}_z[y(\mathbf{x}, \mathbf{z})] = \hat{\mathbf{l}}'(\mathbf{x})\mathbf{V}\hat{\mathbf{l}}(\mathbf{x}) + s^2 \quad (11.22)$$

The result in Equation 11.22 is not an unbiased estimator, because  $E[\hat{\mathbf{I}}'(\mathbf{x})\mathbf{V}\hat{\mathbf{I}}(\mathbf{x})] \neq \mathbf{I}'(\mathbf{x})\mathbf{V}\mathbf{I}(\mathbf{x})$ . In fact,

$$E\{\widehat{\text{Var}}_z[y(\mathbf{x}, \mathbf{z})]\} = \mathbf{I}'(\mathbf{x})\mathbf{V}\mathbf{I}(\mathbf{x}) + \sigma^2\text{tr}(\mathbf{V}\mathbf{C}) + \sigma^2 \quad (11.23)$$

where the matrix  $\mathbf{C}$  is simply  $\text{Var}[\hat{\mathbf{I}}(\mathbf{x})]\sigma^2$ . In other words,  $\mathbf{C}$  is the covariance of the  $\hat{\mathbf{I}}$ 's, apart from  $\sigma^2$ . Recall the role of this matrix in the confidence region on the location of process variance. See Equation 11.19. For standard experimental designs this matrix will be a diagonal matrix, since the  $c_i$ , and  $\hat{\delta}_{jk}$  will all be uncorrelated. However, it should be noted that the diagonal elements, that is, the variances of the  $\hat{\mathbf{I}}'_i(\mathbf{x})$ , depend on  $\mathbf{x}$ . Now, from Equation 11.23 one can work out an unbiased estimator of the process variance, namely,

$$s_z^2[y(\mathbf{x}, \mathbf{z})] = \hat{\mathbf{I}}'(\mathbf{x})\mathbf{V}\hat{\mathbf{I}}(\mathbf{x}) + s^2(1 - \text{tr}\mathbf{V}\mathbf{C}) \quad (11.24)$$

or, for the cases that we have dealing with in this chapter, when the noise variables are uncorrelated with unit variances,

$$s_z^2[y(\mathbf{x}, \mathbf{z})] = \hat{\mathbf{I}}'(\mathbf{x})\hat{\mathbf{I}}(\mathbf{x}) + s^2(1 - \text{tr}\mathbf{C}) \quad (11.25)$$

The estimator in Equation 11.25 will often differ only slightly from the simpler but biased estimator in Equation 11.22. In fact, we used the latter in examples in previous sections. However, there will be occasions when the difference between the two will not be negligible. This is particularly true when the process variance is being computed on or near the design perimeter, because  $\text{tr}\mathbf{C}$  becomes larger as one approaches the design perimeter. As a result, one cannot ignore  $s_z^2[y(\mathbf{x}, \mathbf{z})]$  as an estimator of process variance.

In a comparison of the biased estimator of Equation 11.22 and the unbiased estimator of Equation 11.25, it can be shown that unless the error df in the fitting of the original model in Equation 11.7 is 2 or less, the unbiased estimator in Equation 11.25 is preferable. We will not present the details of that study here. In what follows we discuss the use of the estimated process variance in a ridge-analysis-type procedure for the process variance.

**Ridge Analysis for Process Variance** The reader should recall from Chapter 6 that the purpose of ridge analysis is to produce a locus of points that are constrained optima on the response. This produces a ridge in the design space, which could give rise to future experiments, or could lead to interesting conditions that are of course confined to the experimental region. Applying the same notion to the process variance, we have

$$\min_{\mathbf{x}}\{s_z^2[y(\mathbf{x}, \mathbf{z})]\} \quad \text{subject to} \quad \mathbf{x}'\mathbf{x} = R^2 \quad (11.26)$$

That is, we will develop a locus of points that are each a point of conditionally minimum variance. The use of Lagrange multipliers as in Chapter 6 requires

$$\frac{\partial}{\partial \mathbf{x}} \left[ \hat{\mathbf{I}}'(\mathbf{x})\hat{\mathbf{I}}(\mathbf{x}) + s^2(1 - \text{tr}\mathbf{C}) - \lambda(\mathbf{x}'\mathbf{x} - R^2) \right] \quad (11.27)$$

Now, if the experimental design renders the noise main effects independent of the control  $\times$  noise interactions, and the control  $\times$  noise interaction independent of each other, the matrix  $\mathbf{C}$  reduces to

$$\mathbf{C} = \begin{bmatrix} \mathbf{x}'\mathbf{D}_{11}\mathbf{x} + \text{Var } c_1 & & & \mathbf{0} \\ & \mathbf{x}'\mathbf{D}_{22}\mathbf{x} + \text{Var } c_2 & & \\ & \mathbf{0} & \ddots & \\ & & & \mathbf{x}'\mathbf{D}_{r_z, r_z}\mathbf{x} + \text{Var } c_{r_z} \end{bmatrix}$$

where  $\mathbf{D}_{jj}$  is an  $r_x \times r_x$  matrix containing  $\text{Var}(\hat{\delta}_{ij})/\sigma^2$  on the  $i$ th main diagonal. Here  $\mathbf{x}' = [x_1, x_2, \dots, x_{r_x}]$ . Thus  $\text{tr}(\mathbf{C})$  becomes

$$\text{tr}(\mathbf{C}) = \sum_{j=1}^{r_z} \mathbf{x}'\mathbf{D}_{jj}\mathbf{x} + \sum_{j=1}^{r_z} c_j$$

We will not supply the details in the differentiation of Equation 11.27. However, setting the partial derivatives in Equation 11.27 equal to  $\mathbf{0}$  gives

$$\left[ \hat{\Delta}\hat{\Delta}' - \sigma^2 \sum_{j=1}^{r_z} \mathbf{D}_{jj} - \mu\mathbf{I} \right] \mathbf{x} = -\hat{\Delta}\mathbf{c} \tag{11.28}$$

where  $\mathbf{c}' = [c_1, c_2, \dots, c_{r_z}]$ . To ensure that solutions are points of constrained minimum variance, values of  $\mu$ , smaller than the smallest eigenvalue of  $\hat{\Delta}\hat{\Delta}' - \sigma^2 \sum_{j=1}^{r_z} \mathbf{D}_{jj}$  are to be used in Equation 11.28.

**Example 11.6 Application of Ridge Analysis** Consider a situation in which there are two control variables and three noise variables. The design variables  $x_1, x_2, z_2$ , and  $z_3$  are varied in a 23-run combined array design that is similar to the one used in Example 11.3. It is a variation of a CCD. The design and data are given in Table 11.10.

Note that as before the design is not a complete CCD. The factorial portion is a  $2^{5-1}$ , and there are no axial points in  $z_1, z_2, z_3$ , because we are not fitting quadratic terms in the  $z$ 's. The fitted response model is

$$\begin{aligned} \hat{y} = & 30.382 - 2.925x_1 - 4.136x_2 + 2.855x_1x_2 + 2.596x_1^2 \\ & + 2.175x_2^2 + 2.736z_1 - 2.326z_2 + 2.332z_3 \\ & - 0.278x_1z_1 + 0.893x_1z_2 + 2.574x_1z_3 + 1.999x_2z_1 \\ & - 1.430x_2z_2 + 1.547x_2z_3 \end{aligned}$$

The important derivatives in the  $z_1, z_2$ , and  $z_3$  directions are given by

$$\hat{\mathbf{i}}(\mathbf{x}) = \begin{bmatrix} \hat{l}_1 \\ \hat{l}_2 \\ \hat{l}_3 \end{bmatrix} = \begin{bmatrix} 2.736 - 0.278x_1 + 1.999x_2 \\ -2.326 + 0.893x_1 - 1.430x_2 \\ 2.332 + 2.574x_1 + 1.547x_2 \end{bmatrix}$$

**TABLE 11.10 The Combined Array Design for Example 11.6**

| Observation | $x_1$ | $x_2$ | $z_1$ | $z_2$ | $z_3$ | $y$     |
|-------------|-------|-------|-------|-------|-------|---------|
| 1           | 1     | -1    | 1     | -1    | -1    | 30.0250 |
| 2           | -1    | 1     | -1    | -1    | -1    | 30.0007 |
| 3           | -1    | -1    | 1     | -1    | -1    | 49.8009 |
| 4           | -1    | -1    | -1    | 1     | -1    | 43.4717 |
| 5           | -1    | -1    | -1    | -1    | 1     | 44.1905 |
| 6           | 1     | 1     | 1     | -1    | -1    | 31.3911 |
| 7           | 1     | 1     | -1    | 1     | -1    | 16.0333 |
| 8           | 1     | 1     | -1    | -1    | 1     | 35.3823 |
| 9           | 1     | -1    | 1     | 1     | -1    | 30.3383 |
| 10          | 1     | -1    | 1     | -1    | 1     | 36.3417 |
| 11          | 1     | -1    | -1    | 1     | 1     | 36.1355 |
| 12          | -1    | 1     | 1     | 1     | -1    | 30.1289 |
| 13          | -1    | 1     | 1     | -1    | 1     | 41.3179 |
| 14          | -1    | 1     | -1    | 1     | 1     | 22.7125 |
| 15          | -1    | -1    | 1     | 1     | 1     | 43.2415 |
| 16          | 1     | 1     | 1     | 1     | 1     | 39.1733 |
| 17          | -2    | 0     | 0     | 0     | 0     | 46.1502 |
| 18          | 2     | 0     | 0     | 0     | 0     | 36.0689 |
| 19          | 0     | -2    | 0     | 0     | 0     | 47.3903 |
| 20          | 0     | 2     | 0     | 0     | 0     | 31.4659 |
| 21          | 0     | 0     | 0     | 0     | 0     | 30.8109 |
| 22          | 0     | 0     | 0     | 0     | 0     | 30.7499 |
| 23          | 0     | 0     | 0     | 0     | 0     | 30.9655 |

The method of ridge analysis requires the minimization of  $\hat{\mathbf{1}}\hat{\mathbf{1}} - s^2\text{tr}(\mathbf{C})$  subject to the constraint  $\sum_{i=1}^2 x_i^2 = 2$ . Here we are assuming that the noise variables are uncorrelated with  $\sigma_z^2 = 1.0$  in the coded metric of the noise variables. In this case one needs to solve Equation 11.28 with particular values of  $\mu$ . The matrix  $\sum_{j=1}^{r_z} \mathbf{D}_{jj}$  simple becomes  $\frac{3}{16}\mathbf{I}_2$ , and  $s^2 = 0.920$  from the analysis. As a result,

$$s^2 \sum_{j=1}^2 \mathbf{D}_{jj} = \begin{bmatrix} 0.1725 & 0 \\ 0 & 0.1725 \end{bmatrix}$$

while the matrix  $\hat{\mathbf{A}}$  is the  $2 \times 3$  matrix containing the control  $\times$  noise interaction coefficients in the fitted regression. The eigenvalues of  $\hat{\mathbf{A}}\hat{\mathbf{A}} - s^2 \sum_{j=1}^{r_z} \mathbf{D}_{jj}$  turn out to be 9.995 and 5.594. As a result, we replace  $\mu$  with values smaller than 5.594. It turns out that for  $\mu = -0.486$  the solution to Equation 11.28 falls on a circle of radius  $\sqrt{2}$ . Indeed,  $(-0.0156, -1.41405)$  is the point of constrained minimum variance at radius  $\sqrt{2}$ . Figure 11.21 displays the result of the ridge analysis. All points on the path in the figure are points of constrained minimum process variance. Obviously, future experiments with  $x_1$  near zero and smaller values of  $x_2$  might be considered.

In much of what has preceded we have built mean and variance response surfaces that are generated from a single response surface that is developed from a model with noise and control variables in the same model. We now discuss alternative methods of constructing the variance model.



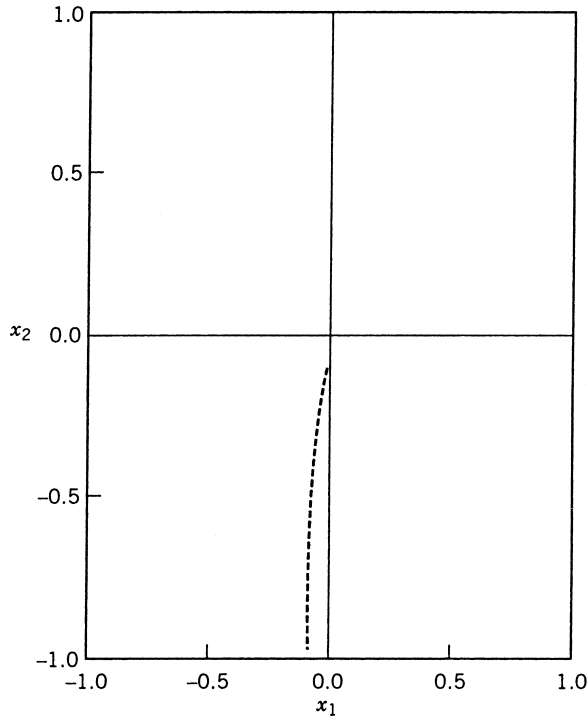


Figure 11.21 Locus of points of minimum constrained process variance for Example 11.6.

11.4.6 Direct Variance Modeling

Bartlett and Kendall (1946) in a classical paper showed that if some design points are replicated, the variance can be modeled directly without serious violation of assumptions by using a **log-linear model** of the

$$\log s_i^2 = \mathbf{x}'_i \boldsymbol{\gamma} + \varepsilon_i^*, \quad i = 1, 2, \dots, d \tag{11.29}$$

where  $s_i^2$  is a sample variance obtained from, say,  $n$  observations taken at each of  $d$  design points. The  $\mathbf{x}'_i \boldsymbol{\gamma}$  portion of the model refers to a linear model in a set of design variables. They demonstrated that if the errors are normal around the mean model; that is, if

$$y_{ij} = \mathbf{x}'_i \boldsymbol{\beta} + \varepsilon_{ij} \sigma_i, \quad i = 1, 2, \dots, d, \quad j = 1, 2, \dots, n$$

where  $\varepsilon_{ij} \sim N(0, 1)$  and  $\ln \sigma_i^2 = \mathbf{x}'_i \boldsymbol{\gamma}$ , then the use of the model in Equation 11.29 leads to approximately normal errors with constant variance. Thus ordinary least squares may be appropriate for modeling variance in a log-linear procedure. Carroll and Ruppert (1988), Engle (1992), Grego (1993), and Vining and Bohn (1998) discuss further details of variance modeling.

Taguchi revived interest in variance modeling. There are many situations where the variance can be directly modeled using the log-linear model approach. These include replicated designs (not all design points have to be replicated) and situations where **repeat**

**measurements** rather than true replicates are taken at the design points, Vining and Schaub (1996) discuss how one might allocate replicate runs in designs so as to obtain a variance model. In Example 3.3 we illustrated how a variance model could be built using duplicate runs. In this situation, the variance reflects **within-run variability**, but usually that is the portion of the total variability that is of most interest to the experimenter. Indeed, a crossed array design also lends itself to variance modeling.

The use of the log-linear model also often results in *simplicity*. The log transformation often reduces the effects of curvature and interaction, thereby making the results easier to interpret.

Once a log-linear variance model is fitted, the mean response surface model in the control variables can be found in the usual way as discussed in Chapter 6. Thus the dual response surface approach for finding robust or optimum conditions can be carried out. Vining and Myers discuss the approach, with illustrations, in their contribution to the panel discussion edited by Nair (1992). Also see Vining and Myers (1990).

**Example 11.7 The Wave Solder Experiment** Consider again Example 11.1 where it was of interest to determine conditions on the control variables that produce a low value of response (solder defects per million joints) but also provide consistency (low variability). The data in Example 11.1 were used to develop two models, one for the mean response (standard linear regression model) and one for the variance via the log-linear variance model described here. Recall that the crossed array contains a  $2_{III}^{5-2}$  inner array design and a  $2_{III}^{3-1}$  in the outer array. Thus, four outer array points can be used for the sample variances. The fitted models, including significant terms, are given by

$$\hat{y} = 197.175 - 28.525x_1 + 57.975x_2$$

for the mean and

$$\ln(s^2) = 7.4155 - 0.2067x_1 - 0.0309x_2$$

for the variance. Here  $x_1$  is the solder temperature and  $x_2$  is the flux density. Figure 11.22 shows the graphical analysis involving both models. The mean and the standard deviations are plotted. The latter was found via the expression

$$s = [e^{7.4155-0.2067x_1-0.0309x_2}]^{1/2}$$

As one would expect from the analysis in Example 11.1, high temperature and low flux density minimize the mean number of errors and the variability simultaneously. However, the graphical dual response surface procedure allows the user to see trade-offs for better understanding of the process. For the example the user may be dissatisfied with the extreme solder temperature and flux density, particularly the temperature. The graphs provide alternative conditions with resulting mean and variance estimates. It would be of interest to see the results when the variance model is computed from the **variance operator** on a model containing both control and noise variables—that is, the procedure discussed in Sections 11.4.3 and 11.4.4. See Exercise 11.25.

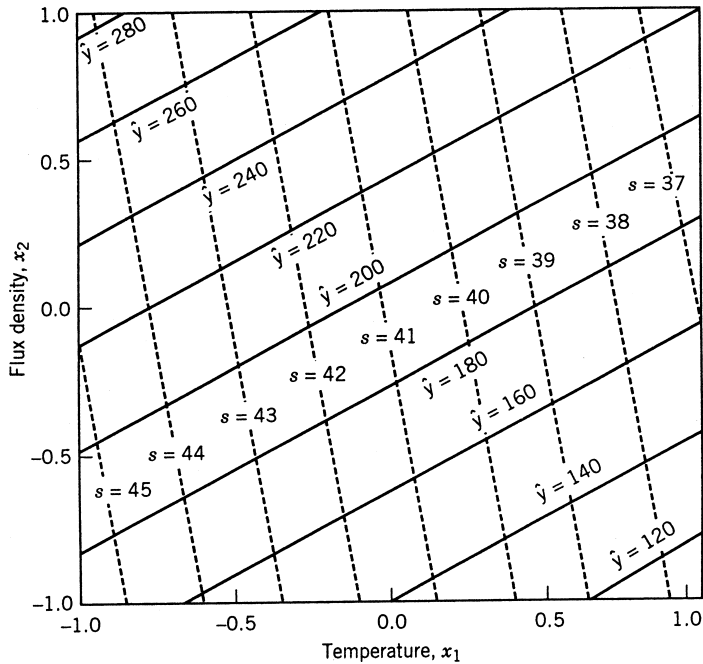


Figure 11.22 Mean and variance models for wave soldering from Example 11.7.

#### 11.4.7 Use of Generalized Linear Models

Much of this chapter has dealt with joint modeling of the mean and variance. These two models can then be used simultaneously or jointly to arrive at candidates for operating conditions. It is instructive to discuss the role of the **generalized linear model** (GLM) in the formulation of the dual RSM procedure. We learned in Chapter 10 that GLMs are an important component in RSM modeling. Often responses are clearly nonnormal, for example, Poisson counts and binary responses. It turns out that even if the basic response is normal, the modeling of variance, either from replication or from nonreplication experiments, is an excellent example where the use of the GLM can be important. The reader will note from what follows that **iterative procedures** can easily be developed that involve two GLMs, one for the mean model and one for the variance model. We begin with the case where either a replicated experiment is made, or noise variables are involved in a crossed array.

**Case of Replicated Experiments or a Crossed Array** Let us assume initially that an experimental design allows for computation of sample variances  $s_i^2$  ( $i = 1, 2, \dots, d$ ) where  $d$  is the number of design points. The  $s_i^2$ -values come either from design replication or from variance information retrieved from an outer array. Earlier in the chapter we considered the modeling of the response  $\ln(s_i^2)$  for the **variance model**. The log transformation results in approximate homogeneity of variance, and hence ordinary least squares was considered in conjunction with a simple one-stage model development. Indeed, if the mean is to be modeled simultaneously with the variance, **weighted least squares** is an appropriate technique. Note, however, that in this approach the distribution of the  $s_i^2$  was not taken into

account. As a result the estimation procedure described in Section 11.4.6 clearly fails to exploit the distribution of the sample variance.

The random variables  $s_1^2, s_2^2, \dots, s_d^2$  are independent estimates of the variances at each design point. Consider now a typical value say  $s_i^2$ . If the responses  $y_{i1}, y_{i2}, \dots, y_{in}$  ( $n$  is the sample size at each design point) are normal, then

$$\frac{s_i^2(n-1)}{\sigma_i^2} \sim \chi_{n-1}^2$$

The  $\chi_{\nu}^2$  distribution is a special case of a **gamma distribution** with dispersion parameter  $2/\nu$ . Recall that the mean and variance of  $\chi_{\nu}^2$  are  $\nu$  and  $2\nu$  respectively. Thus a GLM fit for the variance model with a gamma distribution and known dispersion parameter is indicated. In addition, there are several attractive properties of the log-linear model for  $s_i^2$ , so a **log link** is certainly appropriate. As a result, an efficient procedure for the development of the models is as follows.

1. Use a gamma GLM on the  $s_i^2$ -responses with log link and dispersion parameter  $2/\nu$  to fit the variance model

$$\ln \sigma_i^2 = \mathbf{u}'_i \boldsymbol{\gamma} + \varepsilon_i, \quad i = 1, 2, \dots, d$$

Here we use  $\mathbf{u}'_i \boldsymbol{\gamma}$  as opposed to  $\mathbf{x}'_i \boldsymbol{\beta}$  to allow for possible differences between mean model terms and variance model terms.

2. Use  $v_i = e^{\mathbf{u}'_i \hat{\boldsymbol{\gamma}}}$  as weights to compute  $\mathbf{V} = \text{diag}\{v_1, v_2, \dots, v_d\}$ . Here the  $\hat{\boldsymbol{\gamma}}$  come from the estimator in step 1.
3. Use  $\mathbf{V}$  as a weight matrix to fit the mean model

$$E(y_i) = \mathbf{x}'_i \boldsymbol{\beta}, \quad i = 1, 2, \dots, d$$

with  $\boldsymbol{\beta}$  estimated by  $\mathbf{b} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{V}^{-1}\mathbf{y}$  as a weighted least squares estimator.

It should be understood that the matrix  $\mathbf{X}$  contains model terms for the mean, and the model in Equation 11.29 can in fact contain a subset of these terms.

**Case of Unreplicated Experiments** For unreplicated experiments more complications arise in the case of the variance model, as was pointed out in earlier discussions in this chapter. In fact, as before, it should be noted that the quality of the variance model is very much dependent on a proper choice of the model for the mean, since the source of variability is the **squared residual**  $e_i^2 = (y_i - \mathbf{x}'_i \mathbf{b})^2$  for  $i = 1, 2, \dots, n$  design runs.

For the fundamental variance model we will still deal with the log-linear structure (Eq. 11.29). In order to motivate the use of the GLM for this case, consider the unobserved model error  $\varepsilon_i = y_i - \mathbf{x}'_i \boldsymbol{\beta}$ . As a starting point, we assume that

$$\sigma_i^2 = E(\varepsilon_i^2) = e^{\mathbf{u}'_i \boldsymbol{\gamma}} \tag{11.30}$$

where  $\varepsilon_i^2 \sim \sigma_i^2 x_i^2$ .

**Maximum Likelihood Estimation of  $\beta$  and  $\gamma$  Using GLM** In the dual modeling procedure for unreplicated experiments it is helpful to achieve efficiency of joint estimation of the mean model coefficients  $\beta$  and the variance model coefficients  $\gamma$  through maximum likelihood with the GLM as the analytical tool. We can write the mean model as  $\mathbf{y} = \mathbf{X}\beta + \varepsilon$  with  $\text{Var}(\varepsilon) = \mathbf{V}_{n \times n}$  and again  $v_i = e^{u_i \gamma}$ . Now, the MLE for  $\beta$  is merely weighted least squares,

$$\mathbf{b} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$$

If we use the random vector  $\mathbf{s}' = (\varepsilon_1^2, \varepsilon_2^2, \dots, \varepsilon_n^2)$ , we have a set of independent  $\chi_1^2$  random variables with the properties shown in Equation 11.30. It is interesting, though not surprising, that if we consider maximum likelihood estimation for  $\beta$  and  $\gamma$ , two separate algorithms emerge. However, the maximum likelihood estimator of  $\beta$  involves  $\gamma$  through the matrix  $\mathbf{V}$ , and the maximum likelihood estimator of  $\gamma$  clearly involves  $\beta$ , since the data in the variance model, namely the  $\varepsilon_i$ , involve  $\beta$ . As a result, an iterative procedure is required. We will not write the likelihood function here, but details appear in Aitken (1987). Aitken also pointed out the simplicity provided by the fact that computation of the maximum likelihood estimator for both  $\beta$  and  $\gamma$  can be done with GLMs with an identity link and a normal distribution on the mean model for estimation of  $\beta$  and a log link with a gamma distribution on the variance model for estimation of  $\gamma$ . Indeed, these two procedures can be combined into an iterative procedure. The dispersion parameter for the gamma link is again equal to 2. We must use the **squared residuals**  $e_i^2 = (y_i - \mathbf{x}'_i\mathbf{b})^2$  as the data for the gamma GLM, and raw  $y_i$ -values as data for the normal GLM. The method follows:

1. Use OLS to obtain  $\mathbf{b}_0$  for the mean model  $y_i = \mathbf{x}'_i\beta + \varepsilon_i$ .
2. Use  $\mathbf{b}_0$  to compute  $n$  residuals  $e_i = y_i - \mathbf{x}'_i\mathbf{b}_0$  for  $i = 1, 2, \dots, n$ .
3. Use the  $e_i^2$  as data to fit the variance model with regressors  $\mathbf{u}$  and a log link with dispersion parameter 2. This is done via IRLS with GLM technology.
4. Use parameter estimates from step 3, namely the  $\hat{\gamma}_i$ , to form the matrix  $\mathbf{V}$ .
5. Use  $\mathbf{V}$  with weighted least squares to update  $\mathbf{b}_0$  to  $\mathbf{b}_1$ .
6. Go back to step 2 with  $\mathbf{b}_1$ , replacing  $\mathbf{b}_0$ .
7. Continue to convergence.

The entire algorithm for the model is not difficult to construct and is a useful alternative for dual estimation of the mean and variance models. However, it has the disadvantage that the estimation of the parameters in the variance model will be biased because the maximum likelihood estimation of  $\gamma$  does not allow for estimation of  $\beta$  in the formation of the residuals. A somewhat different procedure serves as an alternative. This procedure makes use of **restricted maximum likelihood** (REML), which was discussed briefly in Chapter 10. The REML procedure does adjust nicely for the bias due to estimation of  $\beta$  in the squared residuals. In addition, this procedure can also be put into a GLM framework.

**Restricted Maximum Likelihood for Dual Modeling** The source of the possible difficulties with ordinary maximum likelihood estimation as discussed above is that the response data for the variance model, namely the  $e_i^2$ , do not truly reflect  $\sigma^2$  unless the estimator  $\mathbf{b}$  is the true parameter vector  $\beta$ . As a result, one cannot expect the estimator of  $\gamma$  from the procedure to be unbiased. Of course, the amount of bias may be quite small. We will deal

with this subsequently. However the IRLS procedure developed for REML is quite simple and should be considered by the practitioner.

Consider now  $E(e_i^2)$  for  $i = 1, 2, \dots, n$ . We know that for a weighted least squares procedure with weight matrix,

$$\begin{aligned} \text{Var}(\mathbf{e}) &= \text{Var} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = \text{Var} [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}] \mathbf{y} \\ &= [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}] \mathbf{V} [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}]' \\ &= \mathbf{V} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}' \end{aligned} \tag{11.31}$$

Since  $E(e_i) = 0$ , then  $E(e_i^2) = \text{Var}(e_i)$ . The matrix  $\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'$  may seem like a **hat matrix** as defined in ordinary multiple regression. However, the matrix that is most like a hat matrix for weighted regression is

$$\mathbf{H} = \mathbf{V}^{-1/2}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1/2} \tag{11.32}$$

The matrix  $\mathbf{H}$  in Equation 11.32 is indempotent and plays the same role as  $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$  does for standard nonweighted (OLS) regression. Now,  $\mathbf{H}$  has the same diagonal elements as  $\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}$ . Thus the diagonal elements in  $\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'$  from Equation 11.31 are also the diagonal elements of  $\mathbf{H}\mathbf{V}$ . As a result, if we consider only the diagonal element involvement in Equation 11.31 we have

$$E(e_i^2) = \sigma_i^2 - h_{ii}\sigma_i^2. \tag{11.33}$$

where  $h_{ii}$  is the  $i$ th hat diagonal, that is, the  $i$ th diagonal of  $\mathbf{H}$  in Equation (11.32).

From Equation 11.33 the adjustment on the response is to use  $e_i^2 + h_{ii}\hat{\sigma}_i^2$  rather than  $e_i^2$ . Here the quantity  $h_{ii}\hat{\sigma}_i^2$  is a **bias correction**. The estimation of  $\beta$  and  $\gamma$  are interwoven and thus an iterative scheme is used, once again with the aid of GLM. The iterative procedure is as follows:

1. Calculate  $\mathbf{b}_0$  from OLS.
2. Calculate the  $e_i$  and a set of responses  $z_{0i} = e_i^2 + h_{ii}s^2$ , where  $h_{ii}$  are the OLS hat diagonals from  $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ , and  $s^2$  is the error mean square.
3. Use the  $z_{0i}$  ( $i = 1, 2, \dots, n$ ) as responses in GLM with a gamma distribution with dispersion parameter 2 and log link to calculate the initial  $\hat{\gamma}$ , say  $\hat{\gamma}_0$ .
4. Use the  $\hat{\gamma}_0$  to produce weights from  $\hat{\sigma}_i^2 = e^{u_i\hat{\gamma}_0}$ , and calculate the weight matrix  $\mathbf{V}_0$ . Calculate a new estimate of the vector  $\beta$  as  $\mathbf{b}_1 = (\mathbf{X}'\mathbf{V}_0^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_0^{-1}\mathbf{y}$ .
5. Calculate new residuals from  $\mathbf{b}_1$  and thus new  $z_i = e_i^2 + h_{ii}\hat{\sigma}_i^2$ , where  $h_{ii}$  is the  $i$ th hat diagonal from the matrix as written in Equation 11.32. Here  $\mathbf{V}$  is diagonal with the current values  $\hat{\sigma}_i^2$  on the main diagonal.
6. Use the new  $z_i$  as responses with a gamma GLM and log link to calculate a new  $\hat{\gamma}$ . Return to step 4 with  $\hat{\gamma}$  replacing  $\gamma_0$ .

The procedure continues with variance fitted values being used for weights in the development of the mean model. The mean model continues to provide the residuals from which the  $z_i$  are developed. Upon convergence, both mean and variance models emerge.

Note the close resemblance between the iterative procedure here and the maximum likelihood estimation algorithm in the previous section. The structure is the same except that at each iteration the gamma GLM is conducted with an **adjusted response**, the adjustment coming via  $h_{ii}\hat{\sigma}_i^2$ , where  $\hat{\sigma}_i^2$  is the estimator of  $\sigma_i^2$  from the previous iteration. The procedure does maximize a **restricted log likelihood**, which is not shown here. For more details the reader should see Engle and Huele (1996). For computational details, consult Harvey (1976). For further discussion of joint modeling with two interwoven GLMs see Nelder and Lee (1991) and Lee and Nelder (1998).

Obviously the correction for bias introduced through REML may be very helpful. However, for large experiments in which  $n \gg p$  where  $p$  is the number of parameters, the bias adjustment will often be negligible.

## 11.5 EXPERIMENTAL DESIGNS FOR RPD AND PROCESS ROBUSTNESS STUDIES

Taguchi advocated the use of crossed arrays (or product arrays) in which separate designs are used for the control and noise variables, and the final design is the cartesian product of the inner array (control array) and the outer array (noise array). Consequently, crossed array designs are sometimes called **product array** designs. The impetus for the design is the SNR, which is used to provide a summary statistic for the data in the outer array. In the same vein,  $\ln s^2$ , discussed in the previous section, is a basic response for variance modeling that can be employed in the crossed array design. However, we have also indicated how, in many instances, the crossed array design can be very costly, with degrees of freedom for interaction among the control factors often being sacrificed in favor of large numbers of degrees of freedom for control  $\times$  noise interactions. If robust parameter design and process robustness studies are to be accomplished, some control  $\times$  noise interaction must be exploited. However, a more efficient type of design is needed.

### 11.5.1 Combined Array Designs

The crossed array was clearly born out of the necessity of having considerable information on which to base the SNR. However, if one focuses on the **response model** approach in which a single model is constructed for both  $\mathbf{x}$  and  $\mathbf{z}$ , with the concomitant development of a mean response surface via  $E_{\mathbf{z}}[y(\mathbf{x}, \mathbf{z})]$  and a variance response surface via  $\text{Var}_{\mathbf{z}}[y(\mathbf{x}, \mathbf{z})]$ , then the crossed array is not needed. For this approach, the **combined array**, in which a design is chosen to allow estimability of a reasonable model in  $\mathbf{x}$  and  $\mathbf{z}$ , is often less costly and quite sufficient. The word “combined” implies that separate arrays are not used, but rather one array that considers a combined  $\mathbf{x}$  and  $\mathbf{z}$  is the appropriate design. We have already seen the use of combined array designs in Examples 11.3, 11.4, and 11.6. Combined array designs have been suggested by many authors, including Welch et al. (1990), Shoemaker et al. (1991), Montgomery (1991), Lucas (1989, 1994), Myers, Khuri, and Vining (1992), and Box and Jones (1989).

Combined arrays offer considerably more flexibility in the estimation of effects or regression coefficients as well as savings in run size. The designs that are chosen generally

are those that offer the concept of **mixed resolution** [see Lucas (1989)]. That is, because control  $\times$  noise interactions are so vital in modeling as well as diagnostics, the resolution of the design for these interactions is higher than that for the noise  $\times$  noise interactions, which are generally considered not as important. The resolution for the control  $\times$  control interactions may also be higher if they are known to be important. In what follows we consider situations in which models are first-order (linear effects) or first-order plus interaction in the control variables.

**Example 11.8 A Mixed Resolution Design** This example, from Shoemaker et al. (1991) shows the superiority of the combined array. Suppose there are three control factors  $x_1, x_2, x_3$  and three noise factors  $z_1, z_2, z_3$ . An obvious crossed array is as follows:

$$2_{\text{III}}^{3-1} \times 2_{\text{III}}^{3-1} \quad (\text{design 1})$$

giving a total of 16 runs. This crossed array could be viewed as a  $2^{6-2}$  combined array with defining relations  $I = x_1x_2x_3 = z_1z_2z_3 = x_1x_2x_3z_1z_2z_3$ . Clearly no interactions in the control factors can be estimated. Six main effects are estimable only if no two factor interactions are important. In addition to six main effects, nine additional degrees of freedom are available, and they are all control  $\times$  noise interactions. These interactions are estimated at the expense of control  $\times$  control interactions.

As an alternative to design 1, a combined array offers considerably more flexibility. One such design would be a  $2^{6-2}$  with defining relations

$$I = x_1x_2x_3z_1 = z_1z_2z_3 = x_1x_2x_3z_2z_3 \quad (\text{design 2})$$

This design illustrates nicely the concept of mixed resolution. The design is resolution III with regard to noise  $\times$  noise interactions and resolution IV with regard to other interactions. Indeed, the three control  $\times$  control interactions can be estimated if one is willing to assume that  $x_1z_1, x_2z_1,$  and  $x_3z_1$  are negligible. In other words, one has the option to trade control  $\times$  noise interactions for control  $\times$  control interactions if appropriate assumptions are met.

**Example 11.9 A Design for Four Control and Two Noise Factors** Consider a situation in which there are four control factors  $x_1, x_2, x_3,$  and  $x_4$  and two noise factors  $z_1$  and  $z_2$ . Assume it is known that  $x_1x_4, x_2x_4,$  and  $x_3x_4$  are important. One potential crossed array design is the 32-run design

$$2_{\text{IV}}^{4-1} \times 2^2 \quad (\text{design 1})$$

The 31 available degrees of freedom include all six main effects and 12 two-factor interactions  $x_1x_2, x_1x_3, x_1x_4, z_1z_2, x_1z_1, x_2z_1, x_3z_1, x_4z_1, x_1z_2, x_2z_2, x_3z_2,$  and  $x_4z_2$ . This represents 18 degrees of freedom. This means that 13 degrees of freedom are being used to estimate higher-order control  $\times$  noise interactions. On the other hand, a combined array  $2_{\text{VI}}^6 - 1$  with 32 runs using

$$I = x_1x_2x_3x_4z_1z_2 \quad (\text{design 2})$$

is much more appropriate. This design allows estimation of all six main effects and all 15 two-factor interactions. Thus three higher-order interactions in design 1 are being exchanged for three control  $\times$  control interactions, which will likely be more important.



An alternative combined array for this situation can be found using **computer-generated design algorithms**. In this case the model contains six main effects and 12 two-factor interactions. There are many designs with less than 32 runs available. Here we show a 22-run design from Shoemaker et al. (1991):

$$\mathbf{D} = \begin{matrix} & x_1 & x_2 & x_3 & x_4 & z_1 & z_2 \\ \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 \\ -1 & -1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & -1 & 1 & -1 \\ -1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & 1 \\ -1 & -1 & 1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 & 1 \\ -1 & 1 & -1 & -1 & -1 & 1 \\ -1 & -1 & 1 & 1 & -1 & 1 \end{bmatrix} \end{matrix}$$

This design allows efficient estimation of all important terms and assigns three degrees of freedom for lack of fit.

### 11.5.2 Second-Order Designs

The flexibility of the combined array and the use of composite designs and computer-generated designs extend nicely to the use of models that are second-order in the control variables and even second-order interactions in the noise variables. Note that the pure quadratic terms for the noise factors are typically not needed, and hence are omitted from the model. Here the mixed resolution idea for the CCD is very important. Consider, for example, a crossed array in the case of three control variables and two noise variables. The control array may be a  $3^{3-1}$  and the noise array a  $2^2$  (36 total runs) to accommodate



As a result, 40 observations plus center runs are used to estimate 30 model terms. Note the careful selection of the defining relations. By definition, the factorial portion is resolution IV. However, it is of higher resolution in the control variables, because no two  $x_i x_j$  interactions are aliased. In addition, it is effectively a higher-resolution design for the control  $\times$  noise interactions, because no two-factor control  $\times$  noise interactions are aliased with each other, nor are they aliased with two-factor interactions in the control variables.

### 11.5.3 Other Aspects of Design

Because the response surface approach puts so much emphasis on efficient estimation of the appropriate response model in  $\mathbf{x}$  and  $\mathbf{z}$ , with flexibility needed in the model terms selected, the combined array is altogether appropriate. In fact, the response model approach and the combined array are quite compatible. Many times the crossed array will involve an excessive number of experimental runs as a result of higher-order control  $\times$  noise interactions. Borkowski and Lucas (1997) discuss mixed resolution combined arrays in detail and give a catalog of designs. Borror and Montgomery (2000) present an application of a mixed resolution design and compare its performance with that of a conventional crossed array design. Borror (1998) is an extensive study of crossed array designs for the robust design problem. She suggests that optimal designs are very appropriate as alternatives to modified versions of standard RSM designs. Borror et al. (2002) develop methodology to evaluate the predictive performance of combined array designs for both the mean model and the slope of the response model (recall the critical role played by the slope with respect to variability in  $y$  transmitted from the noise variables). They point out that the prediction variance of the response model does not tell the complete story regarding design performance and that both the prediction variance of the mean model and the variance of the slope should be evaluated. Variance dispersion graphs and fraction of design space plots from Chapter 9 are an ideal way to do this.

To develop the scaled prediction error variances necessary to study the mean and variance models obtained from the response model, we first review some details concerning least squares fitting of the response model. In general, we may write the response model in the general linear model form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where the matrix  $\mathbf{X}$  contains  $1 + 2r_x + r_x(r_x - 1) + r_z + r_x r_z$  columns representing the intercept,  $\beta_0$ , the linear effects and pure quadratic effects of the  $r_x$  control variables, the two-factor interactions among the control variables, the main effects of the  $r_z$  noise variables, and the control-by-noise variable interactions, respectively. Similarly, the vector  $\boldsymbol{\beta}$  contains the parameters from the mean model corresponding to these factors (the intercept  $\beta_0$  and the elements of  $\boldsymbol{\beta}$ ,  $\mathbf{B}$ ,  $\boldsymbol{\gamma}$ , and  $\boldsymbol{\Delta}$ ). The usual least squares estimate of  $\boldsymbol{\beta}$  is

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$

The matrix  $\mathbf{C} = (\mathbf{X}'\mathbf{X})^{-1}$  is made up of several important submatrices and can be written as in Fig. 11.23. The submatrices,  $\mathbf{C}^{kl}$  ( $k, l = 1, 2, 3$ ) represent the variance–covariance matrices for various subsets of the model parameters. For example,  $\mathbf{C}^{11}$  represents the variance–covariance matrix for the linear, quadratic, and interactions terms involving only the control variables. The submatrix  $\mathbf{C}^{22}$  represents the variance–covariance matrix for



In Equation 11.35,  $\mathbf{x}$  is a fixed set of levels of the controllable variables,  $\mathbf{z}$  is a vector of random variables, and  $\boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{x}$  is a vector of constants. The actual observed value of  $y$  depends on the random variables  $\mathbf{z}$  and  $\varepsilon$  and the specified values of  $\mathbf{x}$ . Now, when  $\hat{y}$  the predicted value of the response, is computed,  $\mathbf{z}$  is fixed at  $\mathbf{z} = \mathbf{0}$ . Thus, the predicted value varies due to the inherent random variability due to the error  $\varepsilon$  and the variability associated with parameter estimation.

A scaled prediction error variance is an appropriate measure to employ in making comparisons among competing experimental designs. The scaled prediction error variance is found by multiplying an expanded form of Equation 11.35 by the number of runs in the design ( $N$ ) and dividing by  $\sigma^2$  to give

$$\frac{N \operatorname{Var}_{\mathbf{z}, \varepsilon} [y(\mathbf{x}, \mathbf{z}) - \mathbf{x}^{(m)'} \hat{\boldsymbol{\beta}}^*]}{\sigma^2} = N \left[ \mathbf{x}^{(m)'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^{(m)} \right] + Nr_2 k_1^2 + 2Nr_2 k_1 k_2 \mathbf{x}' \mathbf{1} + Nr_2 k_2^2 \mathbf{x}' \mathbf{J} \mathbf{x} \tag{11.36}$$

where  $\mathbf{J}$  is a matrix of 1's. Equation 11.36 is in a form that is appropriate for creating VDGs for design evaluation and comparison purposes. The constants  $k_1$  and  $k_2$  represent the contribution of the noise variables and the control-by-noise interactions, respectively. Furthermore, if there are no noise variables in the problem, then  $k_1 = k_2 = 0$ , and Equation 11.36 reduces to the scaled prediction error variance expression that is usually plotted on a VDG.

The variance model assuming  $\operatorname{Var}(\mathbf{z}) = \sigma^2 \mathbf{I}$  is

$$\operatorname{Var}_{\mathbf{z}, \varepsilon} [y(\mathbf{x}, \mathbf{z})] = \sigma_z^2 (\boldsymbol{\gamma}' + \mathbf{x}' \boldsymbol{\Delta}) \mathbf{V} (\boldsymbol{\gamma}' + \mathbf{x}' \boldsymbol{\Delta})' + \sigma^2 \tag{11.37}$$

The vector  $\boldsymbol{\gamma}' + \mathbf{x}' \boldsymbol{\Delta}$  is a vector of partial derivatives of the response surface with respect to the noise variables  $\mathbf{z}$  and is the slope of the response surface in the direction of the noise variables. Thus, the variance of the slope is a direct measure of the precision of estimation associated with the variance model in Equation 11.37. Furthermore, the slope variance is more directly interpretable than a variance expression for Equation 11.37 because it is in the same units as the variance for the mean model.

A general expression for the variance of the slope in direction  $i$ , ( $i = 1, 2, \dots, r_2$ ) is

$$\operatorname{Var}_{z_i} (\text{slope}) = C_{ii}^{22} + \mathbf{x}' \mathbf{C}_i^{33} \mathbf{x} + 2\mathbf{x}' \mathbf{C}_i^{23}, \quad i = 1, 2, \dots, r_2 \tag{11.38}$$

where

$$\mathbf{C}_1^{23} = \begin{pmatrix} C_{11}^{23} \\ \vdots \\ C_{1r_x}^{23} \end{pmatrix}$$

$$\mathbf{C}_i^{23} = \begin{pmatrix} C_{i,(i-1)r_x+1}^{23} \\ \vdots \\ C_{i,ir_x}^{23} \end{pmatrix} \quad \text{for } i > 1$$

and

$$\mathbf{C}_i^{33} = \begin{pmatrix} C_{(i-1)r_x+1, (i-1)r_x+1}^{33} & \cdots & C_{(i-1)r_x+1, ir_x}^{33} \\ \vdots & \ddots & \vdots \\ C_{ir_x, (i-1)r_x+1}^{33} & \cdots & C_{ir_x, ir_x}^{33} \end{pmatrix} \quad i = 1, 2, \dots, r$$

Therefore, a VDG for the slope would be obtained by plotting the average of the slope variances computed from Equation 11.38 along with the maximum and minimum slope variance as a function of the distance of the point from the center of the design.

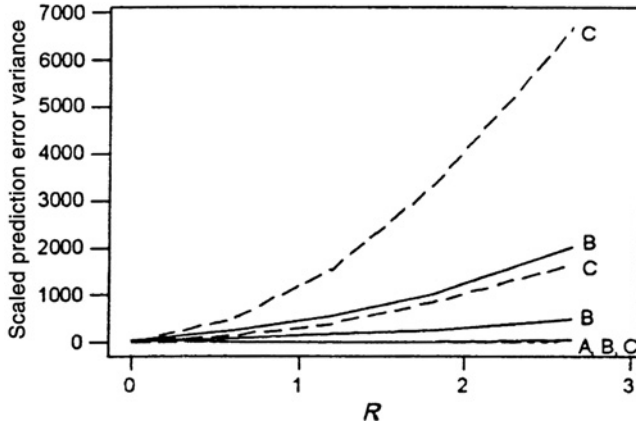
The VDGs constructed for the variance expressed in Equation 11.36 and Equation 11.38 can be used to make comparisons among competing designs for a process robustness study. We now present an illustration of this process.

**Example 11.10** Consider a process robustness study involving  $r_1 = 4$  controllable variables,  $r_2 = 3$  noise variables, and a design region of interest that is spherical. The designs considered by the experimenter are:

1. A “modified” central composite design or CCD ( $N = 75$  runs) with either
  - (a)  $\alpha = 2$
  - (b)  $\alpha = 2.646$
  - or
  - (c)  $\alpha = 2.83$ .
2. A Box–Behnken design (BBD) with  $N = 60$  runs.
3. A modified small composite design or SCD ( $N = 33$  runs) with either
  - (a)  $\alpha = 2$
  - (b)  $\alpha = 2.646$
  - or
  - (c)  $\alpha = 2.83$ .
4. A hybrid design with  $N = 48$  runs.
5. A mixed resolution design (MRD) with  $N = 43$ , and either
  - (a)  $\alpha = 2.646$
  - or
  - (b)  $\alpha = 2.83$ .

Notice that there is considerable disparity in the number of runs for the candidate designs, and the experimenter would like to know if the extra runs in the larger designs are beneficial. The values of  $\alpha$  were chosen based on rotatability and/or spherical properties of either the four-variable (considering only the controllable variables) or the seven-variable (considering all variables) design. The modified CCD and modified SCD are constructed from a standard CCD and SCD, respectively, by removing the axial runs for the noise variables.

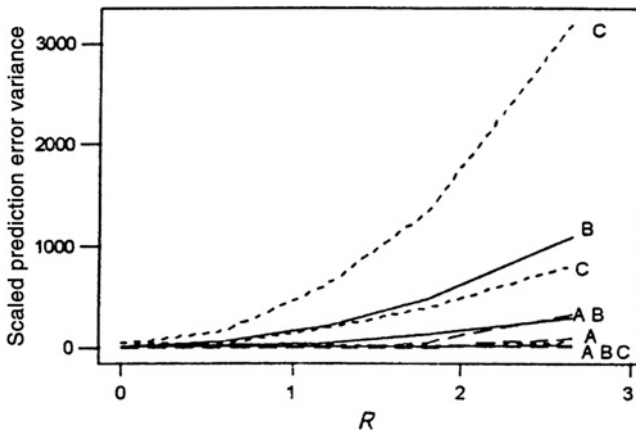
The variance dispersion graphs for the scaled prediction error variance from Equation 11.36 for selected designs with various values of  $k_1$  and  $k_2$  are shown in Figs. 11.24 through 11.26. The minimum, maximum, and average scaled prediction error variance and slope variance at the furthest Euclidean distance from the center for all designs of interest to the



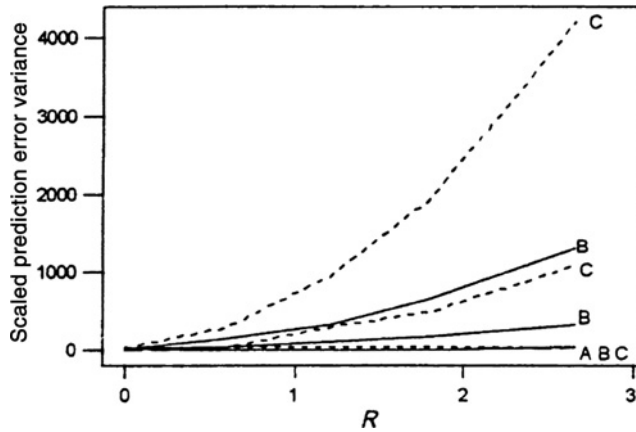
**Figure 11.24** Variance dispersion graph of scaled prediction error variance for modified CCD ( $\alpha = 2.646$ ) with  $k_1 = k_2 = 0$  (A),  $k_1 = k_2 = 0.5$  (B), and  $k_1 = 0.5, k_2 = 1$  (C).

experimenter are provided in Table 11.11. As can be seen from these displays, the prediction error increases as the values of  $k_1$  and  $k_2$  increase. Also, the error increase is affected more by the value of  $k_2$  (the factor pertaining to the control-by-noise interactions) than by the value of  $k_1$ . The variance dispersion graphs for the slope variance for the modified CCD and modified SCD are given in Fig. 11.27. We see that the modified CCD is slope-rotatable regardless of the value of  $\alpha$ ; that is, the slope variance is constant on spheres over the design region. The modified SCD is not slope-rotatable.

There are several important points that should be mentioned about the results in Table 11.11. The modified CCDs and the BBD have 75 experimental runs and 60 experimental runs, respectively. Consequently, these designs may be impractically large to implement. The practitioner would often rely on smaller designs such as the SCD or the hybrid. For four control variables and three noise variables, the modified SCDs have 33 experimental runs, while the hybrid design has 48 experimental runs.



**Figure 11.25** Variance dispersion graph of scaled prediction error variance for modified SCD ( $\alpha = 2.646$ ) with  $k_1 = k_2 = 0$  (A),  $k_1 = k_2 = 0.5$  (B), and  $k_1 = 0.5, k_2 = 1$  (C).

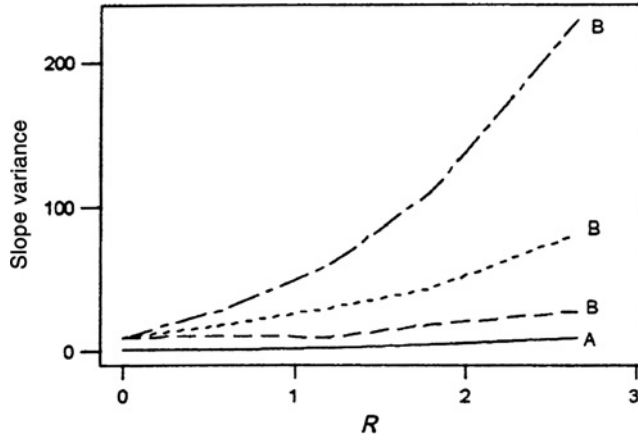


**Figure 11.26** Variance dispersion graph of scaled prediction error variance for hybrid design with  $k_1 = k_2 = 0$  (A),  $k_1 = k_2 = 0.5$  (B), and  $k_1 = 0.5, k_2 = 1$  (C).

**TABLE 11.11** Scaled Prediction Error Variances and Slope Variance for Designs on Spherical Regions with Four Control Variables and Three Noise Variables

| Variance             | Design               | N                   | Scaled Prediction Error ( $k_1, k_2$ ) |            |          | Var(Slope) |
|----------------------|----------------------|---------------------|--|------------|----------|------------|
|                      |                      |                     | (0, 0)                                 | (0.5, 0.5) | (0.5, 1) |            |
| Max                  | CCD $\alpha = 2.83$  | 75                  | 38                                     | 2028       | 6678     | 9.37       |
|                      | $\alpha = 2.646$     | 75                  | 38                                     | 2030       | 6678     | 9.37       |
|                      | $\alpha = 2$         | 75                  | 64                                     | 2052       | 6745     | 9.37       |
|                      | SCD $\alpha = 2.646$ | 33                  | 332                                    | 1089       | 3193     | 230        |
|                      | $\alpha = 2.83$      | 33                  | 306                                    | 1075       | 3179     | 222        |
|                      | BBD                  | 60                  | 204                                    | 1962       | 6157     | 50         |
|                      | Hybrid               | 48                  | 46                                     | 1310       | 4200     | 13.6       |
|                      | MR $\alpha = 2.646$  | 43                  | 33.5                                   | 1110       | 3650     | 10.75      |
|                      | $\alpha = 2.83$      | 43                  | 32                                     | 1116       | 3650     | 10.75      |
|                      | Avg                  | CCD $\alpha = 2.83$ |  | 36         | 484      | 1661       |
| $\alpha = 2.646$     |                      |                     | 36                                     | 484        | 1660     | 9.37       |
| $\alpha = 2$         |                      |                     | 58                                     | 507        | 1684     | 9.37       |
| SCD $\alpha = 2.646$ |                      |                     | 101                                    | 299        | 819      | 82         |
| $\alpha = 2.83$      |                      |                     | 94                                     | 292        | 811      | 79         |
| BBD                  |                      |                     | 204                                    | 564        | 1507     | 50         |
| Hybrid               |                      |                     | 44                                     | 330        | 1080     | 12.2       |
| MR $\alpha = 2.646$  |                      |                     | 32                                     | 290        | 965      | 10.75      |
| $\alpha = 2.83$      |                      |                     | 30                                     | 288        | 965      | 10.75      |
| Min                  |                      | CCD $\alpha = 2.83$ |  | 35         | 36       | 40         |
|                      | $\alpha = 2.646$     |                     | 35                                     | 36         | 39       | 9.37       |
|                      | $\alpha = 2$         |                     | 52                                     | 55         | 58       | 9.37       |
|                      | SCD $\alpha = 2.646$ |                     | 29                                     | 43         | 68       | 28         |
|                      | $\alpha = 2.83$      |                     | 26                                     | 39         | 64       | 28         |
|                      | BBD                  |                     | 204                                    | 221        | 209      | 50         |
|                      | Hybrid               |                     | 42                                     | 45         | 50       | 11.5       |
|                      | MR $\alpha = 2.646$  |                     | 30                                     | 31         | 31       | 10.75      |
|                      | $\alpha = 2.83$      |                     | 26                                     | 27         | 27.5     | 10.75      |

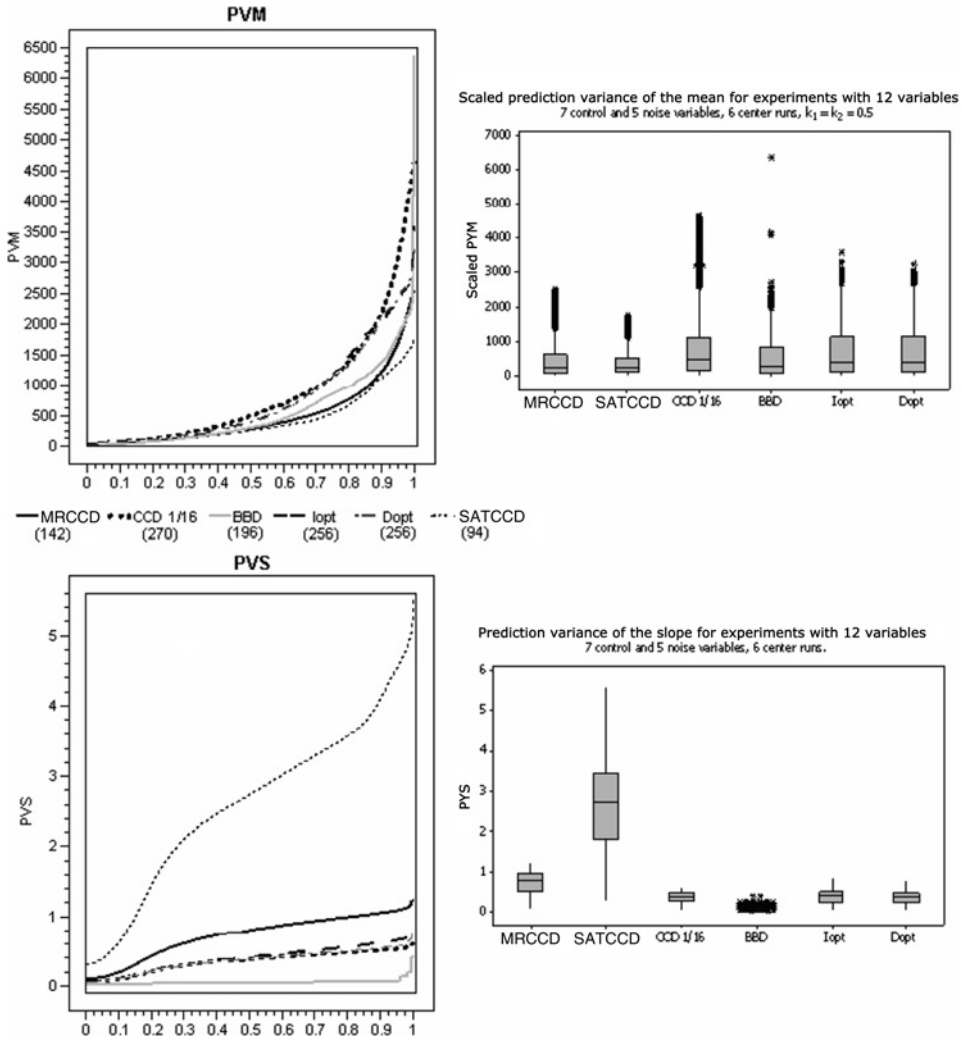




**Figure 11.27** Slope variance for modified CCD (A) and modified SCD (B).

Table 11.11 shows that for the situation where the noise variables do not influence the response ( $k_1 = 0, k_2 = 0$ ), the modified SCD performance is very poor compared with the modified CCD, hybrid, or mixed resolution designs. However, notice that when the noise variables are introduced into the model, the modified SCD is highly competitive with respect to scaled prediction error variance when compared with the modified CCD and the hybrid design. Because the BBD does not estimate the interaction terms as precisely as other designs, it does not perform well overall and would probably not be a good choice. Therefore, it appears that the modified SCD would be a reasonable design when run size is taken into consideration, since there is a considerable difference between  $N = 33$  for the SCDs and  $N = 75$  for the CCDs. However, before deciding that the modified SCD is the best design to use, the slope variances should be investigated. When the slope variances are taken into consideration, the modified SCD is outperformed by all other designs. Consequently, examining both the prediction error variance and the slope variance indicates that the modified CCDs, modified SCDs, hybrid, and BBD are all relatively poor performers with respect to the mean variance, slope variance, or both. In addition, we know that the small composite designs have lower  $G$ - and  $D$ -efficiencies than the other designs, and we recommend against using them. The mixed resolution design outperforms all other designs under consideration with respect to the prediction error variance, the slope variance, and the number of experimental runs. It would clearly be the recommended design in this situation if the experimenter can afford to use 43 runs.

Fraction of design space plots are also useful in facilitating comparisons of prediction variance for the mean and slope in a robust design setting. To illustrate, suppose that we want to compare several designs for a situation where there are seven control variables and five noise variables. The designs considered are a mixed resolution CCD, a modified CCD based on a one-sixteenth fraction in the cube, a BBD, a  $D$ -optimal design, an  $I$ -optimal design, and a CCD where the cube portion of the design is a saturated resolution V fraction. Figure 11.28 shows FDS plots along with box plots of the prediction variance for the mean (PVM) and the prediction variance for the slope (PVS) for these designs. The number of runs for each design is shown in the figure and all designs have six center points. In this scenario the CCD based on a saturated fraction performs well for the mean model but is disastrous for the slope. The BBD is generally a good performer for the mean, except for a



**Figure 11.28** Fraction of design space plots for the mean and slope for 12 variable designs (7 control variables, 5 noise variables).

few outlying values of the prediction variance for the mean, and it performs very well with respect to the slope. Both optimal designs and the mixed resolution CCD are reasonable design choices.

### 11.6 DISPERSION EFFECTS IN HIGHLY FRACTIONATED DESIGNS

In this chapter we have dealt with the important notion of variance modeling where the variance is derivable from the inability to control noise variables. However, there are circumstances in which the practitioner is unable to measure noise variables or perhaps is not aware of what the noise variables are. At any rate, there is nonhomogeneous variance

from one or more sources, and thus there are **both location** and **dispersion effects**. Detection of dispersion effects, positive or negative, is often vital, particularly if the experiment is a pilot experiment and directions are sought for future experiments. For example, consider a  $2^3$  factorial experiment with factors  $A$ ,  $B$ , and  $C$ . In addition, suppose we wish to move to a different experimental region via **steepest ascent** or some other sequential procedure. Suppose the calculated effects are as follows:

$$\begin{aligned} A: & +7.5 \\ B: & +4.7 \\ C: & -3.5 \end{aligned}$$

Now, these calculated effects are **location effects**; that is, they reflect the change in the mean response as one moves from low to high. In this case, future experiments will involve an increase in  $A$ , an increase in  $B$ , and a decrease in  $C$ . However, suppose an increase in  $A$  results in a profound *increase in variance* in the response. We say that  $A$  is a **dispersion effect** as well as a location effect, and thus we are left with a mean–variance tradeoff if we are to consider moving  $A$  to the high level.

In this section the concept of a dispersion effect has not changed. In previous sections a dispersion effect was discovered in a factor if it had a certain type of interaction with a noise variable. Here we are assuming that there are not necessarily any noise variables in the problem, and yet there is a need to detect the existence of factors that influence dispersion. Plots and numerical computation of dispersion effects are useful.

### 11.6.1 The Use of Residuals

We know that much diagnostic information is gained from the use of **residuals**—that is, the values  $e_i = y_i - \hat{y}_i$ , where  $\hat{y}_i$  is the predicted or fitted value from a regression analysis. Residuals provide considerable information about unexplained variability. For example, when residuals are plotted, one expects to see random scatter around zero. Patterns in residuals suggest that some assumption that has been made is being violated. The two most prominent violations are:

1. Functional form of model mis-specified
2. Nonhomogeneous variance

It is violation 2 that concerns us here. It is of interest to determine if the signals from residual plots reflect a type of pattern that would identify certain factors as dispersion effects.

For example, consider a plot of the residuals against the two levels of a factor  $x_j$  shown in Fig. 11.29a. In this case there is no evidence that factor  $x_j$  is a dispersion effect. On the other hand, suppose the plot of residuals takes the form shown in Fig. 11.29b. Now, while one must take considerable care not to read too much into residual plots for small experiments, there is certainly some evidence in Fig. 11.29b that at the high level of  $x_j$  the model error variance is larger than at the low level. While the fitted model here suggests that all three factors are location effects, it appears that  $x_j$  is a dispersion effect, with  $x_j = -1$  favored for a consistent product; that is, one with less variability in the response.

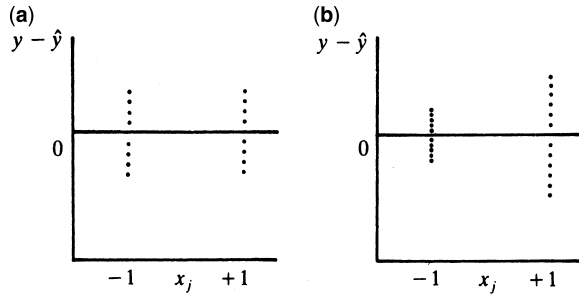


Figure 11.29 Plots of residuals versus factor  $x_j$ .

**Importance of the Fitted Mean Model** It should be understood that the use of residuals, residual plots, and other diagnostic tools for determining dispersion effects is profoundly affected by the fitted mean model, because the residuals used are the residuals from the mean model. If the mean model is not correctly specified, residuals will be clouded by model mis-specification. Thus the detection of dispersion effects depends on the presumption that the fitted model for the mean is a good one. In this regard, the detection of dispersion effects will be most successful when there are only a few well-defined effects to be used in the mean model—that is, when there is **sparsity of effects**.

**11.6.2 Further Diagnostic Information from Residuals**

It should be clear from the illustration in Section 11.6.1 that one is attempting to determine if the product or process variance depends on the design locations, rather than being homogeneous. In the example, we might conclude from the plots that the error variance is constant except when  $x_3 = +1$ . When this type of result is detected, it is hoped that the engineer or scientist has, at least in retrospect, some scientific explanation. However, there still remains the issue surrounding what is meant by a “significant change in variance.” Can we say with assurance that the plots in Section 11.6.1 indicate that  $x_3$  is a dispersion effect? Are there other diagnostic pieces of information apart from the plots?

An important diagnostic for determining if a specific factor, say factor  $i$ , is a dispersion effect is the statistic

$$F_i^* = \ln \left( \frac{s_{i+}^2}{s_{i-}^2} \right) \tag{11.39}$$

See Box and Meyer (1986). Here  $s_{i+}^2$  is the sample variance of the residuals when factor  $i$  is at the high level, and  $s_{i-}^2$  is the sample variance of the residuals when factor  $i$  is at the low level. Obviously a value close to zero is evidence that the spread in the residuals is not significantly different for the two levels, whereas a value large in magnitude (positive or negative) is a signal that the variable has a dispersion effect. A large positive value indicates that  $\sigma_{i+}^2 > \sigma_{i-}^2$ , whereas a large negative effect indicates  $\sigma_{i+}^2 < \sigma_{i-}^2$ . While there should be no intention of using the  $F_i^*$ -value in Equation 11.39 as a formal test statistic, it can be a very useful diagnostic because it has been found that under  $H_0 : \sigma_{i+}^2 = \sigma_{i-}^2$ , the  $F^*$ -values are approximately normally distributed with mean 0 and common variance. This

**TABLE 11.12 Injection Molding Data**

| Standard Order | Actual    |            |            |            |            |                   |                   |                   | Observed Shrinkage (×10) |
|----------------|-----------|------------|------------|------------|------------|-------------------|-------------------|-------------------|--------------------------|
|                | Run Order | $x_1$<br>A | $x_2$<br>B | $x_3$<br>C | $x_4$<br>D | $x_5$<br>E (=ABC) | $x_6$<br>F (=BCD) | $x_7$<br>G (=ACD) |                          |
| 1              | 9         | -          | -          | -          | -          | -                 | -                 | -                 | 6                        |
| 2              | 16        | +          | -          | -          | -          | +                 | -                 | +                 | 10                       |
| 3              | 18        | -          | +          | -          | -          | +                 | +                 | -                 | 32                       |
| 4              | 17        | +          | +          | -          | -          | -                 | +                 | +                 | 60                       |
| 5              | 3         | -          | -          | +          | -          | +                 | +                 | +                 | 4                        |
| 6              | 5         | +          | -          | +          | -          | -                 | +                 | -                 | 15                       |
| 7              | 10        | -          | +          | +          | -          | -                 | -                 | +                 | 26                       |
| 8              | 2         | +          | +          | +          | -          | +                 | -                 | -                 | 60                       |
| 9              | 9         | -          | -          | -          | +          | -                 | +                 | +                 | 8                        |
| 10             | 15        | +          | -          | -          | +          | +                 | +                 | -                 | 12                       |
| 11             | 12        | -          | +          | -          | +          | +                 | -                 | +                 | 34                       |
| 12             | 6         | +          | +          | -          | +          | -                 | -                 | -                 | 60                       |
| 13             | 13        | -          | -          | +          | +          | +                 | -                 | -                 | 16                       |
| 14             | 19        | +          | -          | +          | +          | -                 | -                 | +                 | 5                        |
| 15             | 11        | -          | +          | +          | +          | -                 | +                 | -                 | 37                       |
| 16             | 1         | +          | +          | +          | +          | +                 | +                 | +                 | 52                       |
| 17             | 20        | 0          | 0          | 0          | 0          | 0                 | 0                 | 0                 | 25                       |
| 18             | 4         | 0          | 0          | 0          | 0          | 0                 | 0                 | 0                 | 29                       |
| 19             | 14        | 0          | 0          | 0          | 0          | 0                 | 0                 | 0                 | 24                       |
| 20             | 7         | 0          | 0          | 0          | 0          | 0                 | 0                 | 0                 | 27                       |

may be exploited as a numerical diagnostic, and the  $F_i^*$ -values may be studied via normal probability plots.

**Example 11.11 Dispersion Effects** In an injection molding experiment it is often important to determine conditions in which shrinkage is minimized. In the present experiment [see Montgomery (1991)] seven factors were varied in a  $2^{7-3}$  with four center runs. It was felt that there might be curvature in the system. The factors are  $x_1$  (mold temperature),  $x_2$  (screw speed),  $x_3$  (holding time),  $x_4$  (gate size),  $x_5$  (cycle time),  $x_6$  (moisture content), and  $x_7$  (holding pressure). The defining relations are

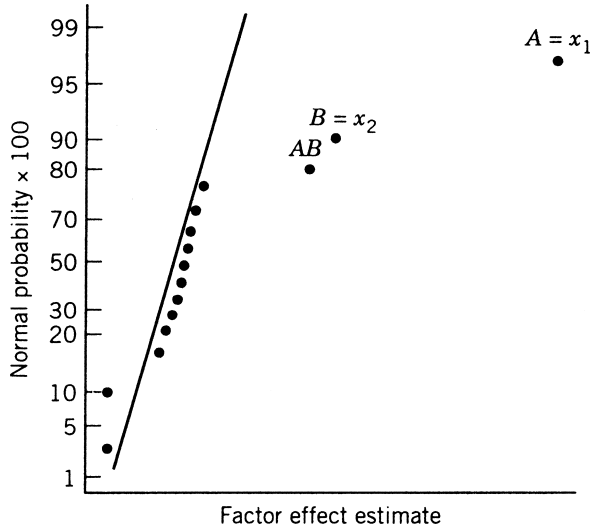
$$x_1x_2x_3x_5 = x_2x_3x_4x_6 = x_1x_3x_4x_7 = I$$

The data appear in Table 11.12. It is of interest to determine settings (within the region of the experiment) that minimize shrinkage. In addition, any dispersion effect that is found must be dealt with, because the variability in the shrinkage must also be taken into account.

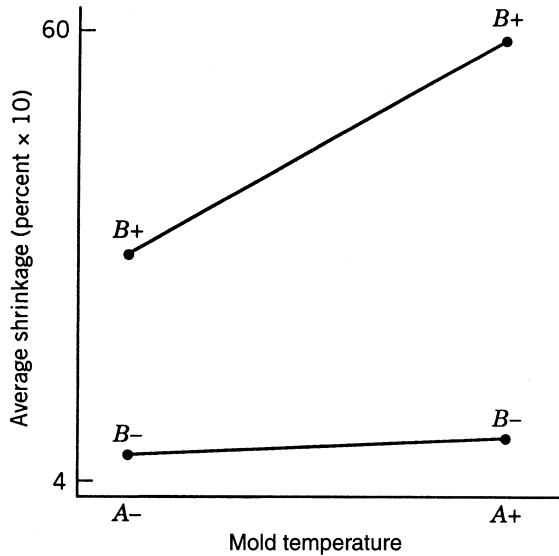
The individual location effects of interest were first placed on a normal probability plot to gain some impression of what model terms are important. Table 11.13 gives the reader a clear indication of what effects are estimable. Fifteen effects, including all main effects and two-factor interactions (and aliases), were plotted on the normal probability plot in Fig. 11.30. The  $x_1$ ,  $x_2$ , and  $x_1x_2$  effects are found to be important from the plot. All three are positive effects. The  $x_1x_2$  interaction plot in Fig. 11.31 is particularly revealing. The use of the screw speed  $x_2$  at the low level minimizes shrinkage regardless of the level of the mold temperature  $x_1$ . In addition, if the mold temperature is difficult to control in the process, the

**TABLE 11.13** Alias Structure for the  $2^{7-3}$  Design in Table 11.12

|   |                |
|---|----------------|
| $A = BCE = DEF = CDG = BFG$               | $AB = CE = FG$ |
| $B = ACE = CDF = DEG = AFG$               | $AC = BE = DG$ |
| $C = ABE = BDF = ADG = EFG$               | $AD = EF = CG$ |
| $D = BCG = AEF = ACG = BEG$               | $AE = BC = DF$ |
| $E = ABC = ADF = BDG = CFG$               | $AF = DE = BG$ |
| $F = BCD = ADE = ABG = CEG$               | $AG = CD = BG$ |
| $G = ACD = BDE = ABF = CEF$               | $BD = CF = EG$ |
| $ABD = CDE = ACF = BEF = BCG = AEG = DEF$ |                |



**Figure 11.30** Normal probability plot of location effects for Example 11.11.



**Figure 11.31**  $AB = x_1, x_2$  interaction plot for Example 11.11.

**TABLE 11.14 Regression and ANOVA for Data of Example 11.11**

| Source of Variation     | Sum of Squares | Degrees of Freedom | Mean Square                | <i>F</i> | Probability > <i>F</i> |
|-------------------------|----------------|--------------------|----------------------------|----------|------------------------|
| Model                   | 6410.687500    | 3                  | 2136.8958333               | 121.645  | 0.0001                 |
| Curvature               | 3.612500       | 1                  | 3.6125000                  | 0.206    | 0.6567                 |
| Error                   | 263.500000     | 15                 | 17.5666667                 |          |                        |
| Residual                | 248.750000     | 12                 | 20.7291667                 |          |                        |
| Pure error              | 14.750000      | 3                  | 4.9166667                  |          |                        |
| Corrected total         | 6677.800000    | 19                 |                            |          |                        |
| Root mean squared error | 4.191261       |                    | <i>R</i> -squared          | 0.9605   |                        |
| Dependent mean          | 27.312500      |                    | Adjusted <i>R</i> -squared | 0.9500   |                        |
| CV                      | 15.35%         |                    |                            |          |                        |

| Variable  | Parameter Estimate | Degrees of Freedom | Sum of Squares | <i>t</i> for <i>H</i> <sub>0</sub> :<br>Parameter = 0 | Probability > [ <i>t</i> ] |
|-----------|--------------------|--------------------|----------------|---|----------------------------|
| Intercept | 27.312500          | 1                  |                |   |                            |
| <i>A</i>  | 6.937500           | 1                  | 7700.062500    | 6.621   | 0.0001                     |
| <i>B</i>  | 17.812500          | 1                  | 5076.562500    | 17.000  | 0.0001                     |
| <i>AB</i> | 5.937500           | 1                  | 564.062500     | 5.667   | 0.0001                     |

−1 level of the screw speed is also a robust choice. So at this point in the analysis it would seem that  $x_2 = -1$  is absolutely necessary.

The next step in the analysis is to consider possible curvature. Table 11.14 summarizes a regression analysis with the model

$$\hat{y} = 27.3125 + 6.9375x_1 + 17.8125x_2 + 5.9375x_1x_2 \tag{11.40}$$

and an analysis of variance that reveals no evidence of curvature. The curvature term in the ANOVA is a result of center points being used in the design (only a single degree of freedom). There is no actual curvature in the system. At this point it would appear that  $x_1 = -1$  and  $x_2 = -1$  are proper settings. It turns out that this would reduce the mean shrinkage by roughly 10%. But what about variability? The regression in Equation 11.40 was used to compute the residuals. The normal probability plot of residuals in Fig. 11.32 do not reveal anything out of the ordinary. However, the residuals were also plotted against levels of the main effects, and the residual plot involving  $x_3$  shown in Fig. 11.33 was particularly enlightening. It seems clear that variability in the residuals is greater at high holding time than at low holding time. The  $s_{(i-)}$  and  $s_{(i+)}$  values as well as the  $F_i^*$  in Equation 11.39 were computed. They appear, along with the residuals themselves, in Table 11.15. Note that the single  $F_i^*$ -value that stands out is that of  $F^* = 2.50$ . To gain more diagnostic insight, a normal probability plot of the  $F^*$ -values appears in Fig. 11.34. Clearly,  $x_3$  stands out. It would not be unreasonable to conclude that  $x_3$  has a significant dispersion effect.

It is of interest now to provide a summary picture of the mean and variability of the shrinkage response. This is particularly easy to do because there appears to be only three

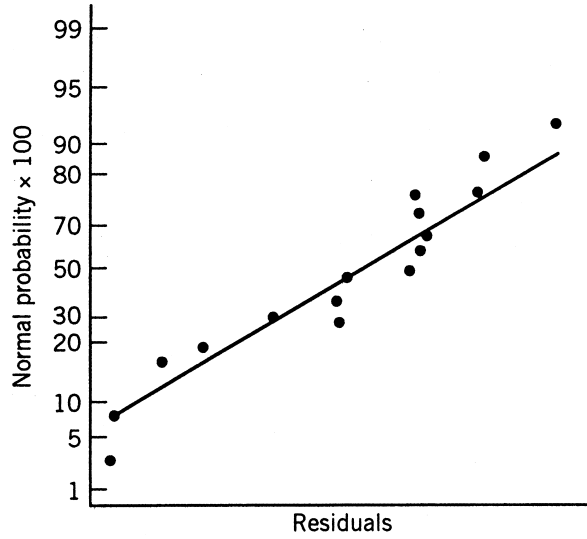


Figure 11.32 Normal probability plot of residuals for Example 11.11.

“main players,”  $x_1$ ,  $x_2$ , and  $x_3$ . The analysis suggests that low  $x_1$ , low  $x_2$ , and low  $x_3$  are the settings that produce, simultaneously, minimum mean shrinkage and minimum variance in shrinkage. That is, the optimum conditions are

$$x_1 = -1, \quad x_2 = -1, \quad x_3 = -1$$

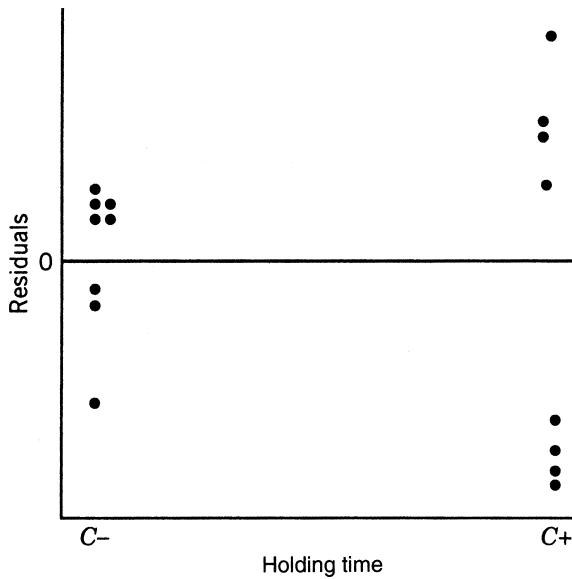


Figure 11.33 Plot of residuals against holding time for Example 11.11.



**TABLE 11.15 Residuals and  $F^*$ -Values for All Effects for Example 11.11**

|         | $x_1$ | $x_2$ | $x_1x_2$ | $x_3$ | $x_1x_3$ | $x_2x_3$ | $x_5$ | $x_4$ | $x_1x_4$ | $x_2x_4$ | $x_1x_2x_4$ | $x_3x_4$ | $x_7$ | $x_6$ | $x_4x_5$ | Residuals |
|---------|-------|-------|----------|-------|----------|----------|-------|-------|----------|----------|-------------|----------|-------|-------|----------|-----------|
| $S_-$   | 4.59  | 4.41  | 4.10     | 1.63  | 4.52     | 4.33     | 4.25  | 3.59  | 2.75     | 4.41     | 3.51        | 3.50     | 3.65  | 3.12  | 3.40     | -2.50     |
| $S_+$   | 3.80  | 4.01  | 4.33     | 5.70  | 3.68     | 3.85     | 4.17  | 4.64  | 3.39     | 4.01     | 4.72        | 3.12     | 3.50  | 3.88  | 4.87     | -0.50     |
| $F_i^*$ | -0.38 | -0.18 | 0.11     | 2.50  | -0.41    | -0.24    | -0.03 | 0.51  | 0.42     | -0.18    | 0.52        | 0.51     | 0.23  | -0.30 | 0.72     | +0.25     |
|         |       |       |          |       |          |          |       |       |          |          |             |          |       |       |          | 2.00      |
|         |       |       |          |       |          |          |       |       |          |          |             |          |       |       |          | -4.50     |
|         |       |       |          |       |          |          |       |       |          |          |             |          |       |       |          | 4.50      |
|         |       |       |          |       |          |          |       |       |          |          |             |          |       |       |          | -6.26     |
|         |       |       |          |       |          |          |       |       |          |          |             |          |       |       |          | 2.00      |
|         |       |       |          |       |          |          |       |       |          |          |             |          |       |       |          | +0.50     |
|         |       |       |          |       |          |          |       |       |          |          |             |          |       |       |          | 1.50      |
|         |       |       |          |       |          |          |       |       |          |          |             |          |       |       |          | 1.75      |
|         |       |       |          |       |          |          |       |       |          |          |             |          |       |       |          | 2.00      |
|         |       |       |          |       |          |          |       |       |          |          |             |          |       |       |          | 7.50      |
|         |       |       |          |       |          |          |       |       |          |          |             |          |       |       |          | -5.50     |
|         |       |       |          |       |          |          |       |       |          |          |             |          |       |       |          | 4.75      |
|         |       |       |          |       |          |          |       |       |          |          |             |          |       |       |          | -6.00     |

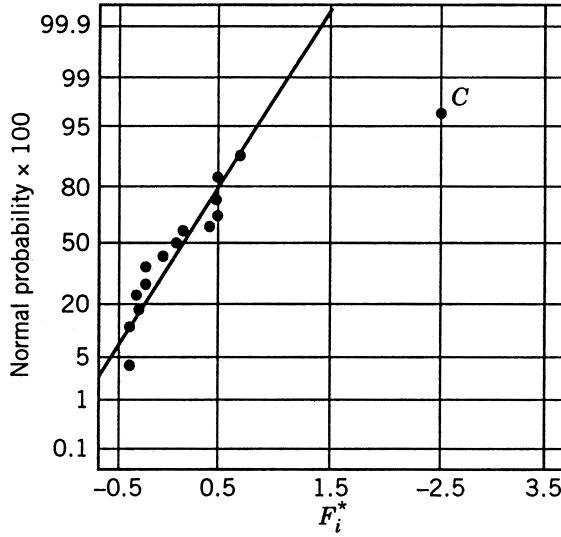


Figure 11.34 Normal probability plots of dispersion effects  $F_i^*$  for Example 11.11.

### 11.6.3 Further Comments Concerning Variance Modeling

In this chapter we have focused on the importance of considering process variability. The major sources of process variability are often noise variables; and when this is the case, the modeling can be handled conveniently through the use of a combined array and of the resulting response surface in  $\mathbf{x}$  and  $\mathbf{z}$  and response surfaces that result for both mean and variance. Another source of variance modeling is the **log-linear model** discussed in Section 11.4.6, where  $\ln(s^2)$  becomes the response. In this context, the  $s^2$  was viewed as the sample variance associated with the outer array in a crossed array, but this need not be the case. The  $s^2$  may be computed from replicated observations or even from repeated measurements on the same experimental unit in, say, a two-level design; in other words, there may be no noise variables in the design and yet a nonconstant variance may be evident through replication ( $n$  runs at each of  $d$  design points) and the use of the model

$$\ln(s_i^2) = f(\mathbf{x}_i) + \varepsilon_i, \quad i = 1, 2, \dots, d \tag{11.41}$$

Additional levels of variance modeling and/or the computation of dispersion effects can be accomplished for cases such as Example 11.11. The experiment contains a large number of factors, and because of cost constraints the design is a highly fractionated two-level design with no replication. Here, formal variance modeling is not necessarily a good practice, due to lack of sufficient information on variability. However, the use of residuals can highlight dispersion effects as in the example. The reader should be aware of the possibility of obtaining variance information and dispersion effects at different levels of experimentation and analysis. In addition, more discussion is needed regarding different methods of calculating dispersion effects.

**What Is a Dispersion Effect? (Case of Replicated Observations)** Box and Meyer (1986) and Nair and Pregibon (1986) discuss the notion of dispersion effects for various types of

scenarios. Consider a situation in which the design is a two-level factorial or fraction and  $n$  observations appear at each data point. The fitted model may be a log-linear model with

$$\ln(s_i^2) = \gamma_0 + \gamma_1 x_{i1} + \gamma_2 x_{i2} + \dots + \gamma_k x_{ik} + \varepsilon_i, \quad i = 1, 2, \dots, d \quad (11.42)$$

In this case, the replication variance is more reliable as process variance information than that obtained from residuals. In the latter case, the reliability of the variance information depends to a great extent on the quality of the mean model that created the residuals.

The least squares estimators of the  $\gamma_i$  in Equation 11.42 are derived from

$$\hat{\boldsymbol{\gamma}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e} \quad (11.43)$$

where  $\mathbf{e}$  is a vector of  $\log s^2$  values. Because the rows of  $\mathbf{X}'$  contain +1 and -1 values depending on whether the specific factor is high or low, and because  $(\mathbf{B}'\mathbf{X})^{-1} = (1/d)\mathbf{I}$ , the parameter estimates for the model 11.42 become

$$\begin{aligned} \hat{\gamma}_0 &= \overline{\ln(s^2)} \\ \hat{\gamma}_i &= \frac{\sum_{j=i+} \ln s_i^2 - \sum_{j=i-} \ln s_j^2}{d}, \quad i = 1, 2, \dots, k \end{aligned} \quad (11.44)$$

Here  $\sum_{j=i+}$  implies that the sum is taken over the  $\ln s^2$  values for which the  $i$ th variable is at the +1 level, while  $\sum_{j=i-}$  implies that sum is over the -1 level. In addition,  $\overline{\ln(s^2)}$  is the average of the  $\ln s^2$  values over the experiment. One could, of course, extend what we know about the relationship between effects and regression coefficients for the mean model. Thus the **computed dispersion effect** for the  $i$ th variable can be written

$$(\text{Dispersion effect})_i = \left[ \overline{\ln(s^2)}_{i+} - \overline{\ln(s^2)}_{i-} \right], \quad i = 1, 2, \dots, k \quad (11.45)$$

where, obviously  $\overline{\ln(s^2)}_{i+}$  is the average  $\ln(s^2)$  value when variable  $i$  is at the high level, and  $\overline{\ln(s^2)}_{i-}$  is the same for variable  $i$  at the lower level. Now, what is done with those dispersion effects or regression coefficients? Normal probability plots can certainly be used to highlight important dispersion effects. The assumption of normality with constant variance for these effects is certainly justified. We suggested earlier that work by Bartlett and Kendall (1946) and others indicates that  $\ln(s^2)$  is approximately normal, and the averaging process strengthens the assumption. Of course, it is quite likely that in most applications where replications are involved, the number of design points and, hence, the number of variables will not be large. As a result, there may not be many effects being studied (including interactions), so the normal probability plot may not contain a very large "sample size." However,  $t$ - or  $F$ -tests in which lack of fit is used as the "error" for testing main effect dispersion effects will certainly supply useful information.

In addition to the dispersion effects given in Equation 11.45, another one, representing a slight variation, has been suggested. Rather than summing the log variances, the alternative dispersion effects takes sums first—that is,

$$\begin{aligned}
 (\text{Dispersion effect}^*)_i &= \ln \left[ \sum_{j=i+} s_j^2 \right] - \ln \left[ \sum_{j=i-} s_j^2 \right] \\
 &= \ln \left[ \frac{\overline{s_{i+}^2}}{s_{i-}^2} \right]
 \end{aligned}
 \tag{11.46}$$

This, of course, is closer in concept to the  $F^*$  dispersion effect discussed earlier for unrepliated experiments.

**The Unrepliated Case** We have already discussed the  $F^*$  dispersion effect calculation and illustrated it with an extensive case study in injection molding. However, it may be of interest to determine the genesis of this dispersion effect, which was developed by Box and Meyer (1986). Consider first an experiment with a two-level unrepliated factorial. Consider the mean model of  $N$  observations,

$$y_i = \mathbf{x}'_i \boldsymbol{\beta} + \varepsilon_i^* \sigma_i, \quad i = 1, 2, \dots, N \tag{11.47}$$

where the  $\varepsilon_i^*$  are independent  $N(0, 1)$  and

$$\ln(\sigma_i^2) = \mathbf{x}'_i \boldsymbol{\gamma} \tag{11.48}$$

Now, clearly the standard deviation changes as  $\mathbf{x}$  moves. The model in Equation 11.48 need not imply that all the variables in the mean model are in the variance model. In fact, the mean model often will contain interactions (or even main effects) not contained in the variance model. We noticed that in our injection molding example.

Consider now the distribution of an error in the mean model—that is, the distribution of

$$\varepsilon_i = y_i - \mathbf{x}_i' \boldsymbol{\beta}$$

We know that

$$\varepsilon_i^2 \sim \sigma_i^2 \chi_1^2$$

and thus

$$\ln \varepsilon_i^2 = \ln \sigma_i^2 + \ln \chi_1^2 \tag{11.49}$$

Thus, one could view Equation 11.49 as a model for the log of the  $i$ th squared residual. Keep in mind however, that we are assuming in Equation 11.49 that *the mean is known*—that is,  $\varepsilon_i$  is not an observed residual  $y_i - \hat{y}_i$ , but rather  $y_i - \mathbf{x}'_i \boldsymbol{\beta}$ . As a result, one can view an *approximate* model of the log squared residual as

$$\ln e_i^2 \cong \ln \sigma_i^2 + \ln \chi_1^2$$

with  $\ln \chi^2$  representing an approximately normal random variable with mean approximately zero and constant variance. From Equation 11.49 we have

$$\ln e_i^2 = \mathbf{x}'_i \boldsymbol{\gamma} + \tilde{\varepsilon}_i \tag{11.50}$$

[it can be shown that  $\text{Var}(\ln \chi^2) \cong 2/\nu$ ], where  $\tilde{\varepsilon}_i$  has approximately mean 0 and variance 2.

At this point one should consider least squares estimation in Equation 11.50. The parameter estimates from which dispersion effects can be are given by

$$\begin{aligned} \tilde{\gamma}_0 &= \overline{\ln e^2} \\ \hat{\gamma}_1 &= \frac{\sum_{j=i+} \ln e_j^2 - \sum_{j=i-} \ln e_j^2}{N} \end{aligned}$$

As a result, an obvious dispersion effect will be  $2\hat{\gamma}_i$ , which is given by

$$(\text{Dispersion effect})_i = \left[ \overline{\ln e^2} \right]_{i+} - \left[ \overline{\ln e^2} \right]_{i-} \tag{11.51}$$

where  $\overline{\ln e^2}$  indicates the average of the log of the squared residual. Again a *normal probability plot*, once a mean model of high quality has been chosen, is certainly a reasonable diagnostic procedure. In fact, dispersion effects will be approximately normal with mean zero (assuming a zero dispersion effect) and constant variance. Thus, any dispersion effect off the straight line is diagnosed as having mean not zero and, thus, being an important dispersion effect.

From the foregoing it should be evident to the reader why  $F_i^*$ , is also a reasonable dispersion effect. As in the case with replicated factorials, the  $F_i^*$ , in Equation 11.39 is much like that in Equation 11.51 except that the  $\sum_j$  operation occurs before taking logs. In this regard, another possible dispersion effect is given by

$$F_i = \frac{\ln \sum_{j=i+} e_j^2 - \ln \sum_{j=i-} \ln e_j^2}{N} = \ln \left[ \frac{\sum_{j=i+} e_j^2}{\sum_{j=i-} e_j^2} \right]$$

Of course this is motivated through the use of the  $\varepsilon_j$  rather than the  $e_j$ . It is reasonable to correct the  $e_j$  for their respective means, so one is using estimates of  $\sigma_{i+}^2$  and  $\sigma_{i-}^2$ . Keep in mind that the existence of a nonzero  $\gamma_i$  in the model of Equation 11.48 presumes that the variances at  $x_i = +1$  and  $x_i = -1$  are distinct. As a result, a useful dispersion effect is given by

$$F_i^* = \ln \left[ \frac{s_{i+}^2}{s_{i-}^2} \right]$$

which was introduced without motivation in Equation 11.39. It is important for the user to understand that the practical use of dispersion effects in unreplicated factorials or fractional

factorials should include diagnostic plots (e.g., normal probability plots) and *not* formal significance testing.

## EXERCISES

- 11.1** Consider Exercise 3.11 in Chapter 3. It is of interest to maximize the etch rate. As a result, it is a larger-the-better scenario. Suppose that  $C$  (gas flow) and  $D$  (power applied to the cathode) are difficult to control in the process.
- Treating  $C$  and  $D$  as noise variables, illustrate, with a drawing, that this design is a crossed array.
  - Analyze the data and use the standard larger-the-better SNR. Do pick-the-winner analysis and thus determine optimum values of  $A$  (anode–cathode gap) and  $B$  (pressure in the reactor chamber).
- 11.2** Consider Example 4.2 in Chapter 4. Suppose, in this process for manufacture of integrated circuits, that the temperature is very difficult to control. There is some concern over variability in wafer resistivity due to uncontrolled variability in temperature. Consider the other factors as control variables.
- From the analysis given in Chapter 4, can any of the control variables be used to exert some influence over this variability? Explain.
  - It is of interest to maximize wafer resistivity and still minimize variability produced by changes in temperature. Can this be done? Explain.
- 11.3** Consider Exercise 11.2.
- Estimate the process variance as a function of  $x_1$ , the implant dose.
  - What assumptions were made in constructing this variance function? Discuss the variance–mean tradeoff here. If we use  $\hat{\mu} - 2\sigma$  as a criterion, give optimal values of implant dose and time.
- 11.4** Consider Exercise 11.3. Because  $E$  (furnace position) and  $D$  (oxide thickness) are unimportant, reduce the design to a  $2^3$  factorial with duplicate runs. Using temperature as a single noise variable, determine the optimal values of implant dose and time. Use a Taguchi analysis with the appropriate SNR ratio for a larger-the-better situation.
- 11.5** Consider Exercise 3.9. Use the concept of dispersion effects to determine if any factors affect the process variance. Comment on your results.
- 11.6** Consider Example 11.1 dealing with solder process optimization.
- Construct and edit a response model involving all control and noise variables.
  - Generate an estimated mean model as functions of the control variables.
  - Generate an estimated process variance function as a function of the control variables.
- 11.7** Consider Example 11.2. The design is criticized for not allowing interaction among the control variables to be studied.
- What design would be a good candidate to replace the crossed array in this example? Do not allow your design to admit a run size any greater than the

design listed in the example. Use a design that allows quadratic effects in the control factors and allows construction of a response model that can be used to generate the process mean and variance models.

(b) Explain why your design is better than that used in the example.

**11.8** Consider the filtration rate data in Example 11.3. Suppose that the filtration rate should be maximized. Suppose, however, that temperature must be treated as a noise variable.

(a) What factors can be used to control the variability in filtration rate due to random fluctuation in temperature in the range  $[-1, +1]$ ?

(b) What levels (high or low) of formaldehyde and stir rate result in minimum process variance? (Study interaction plots.)

**11.9** Consider Exercise 11.8.

(a) Again, consider temperature as a noise variable. Use the fitted regression model as a response model to produce a variance model assuming that one design unit in temperature is  $2\sigma_z$ , where  $\sigma_z$  is the standard deviation of temperature in the process.

(b) What levels of  $C$  and  $D$  result in a minimum process variance?

(c) Discuss the mean–variance trade-off here. Give a quantitative discussion. What are reasonable levels of  $C$  and  $D$ ?

**11.10** Consider a  $2^4$  factorial design. Suppose all six two-factor interaction effects are found to be important. Suppose, in addition, that  $A$  and  $B$  are control variables and  $C$  and  $D$  are noise variables. Suppose  $\sigma_{z_1} = \sigma_{z_2} = 1.0$

(a) Write out the form of the mean model.

(b) Write out the form of the variance model.

**11.11** Consider an experimental situation with  $A(x_1)$ ,  $B(x_2)$ , and  $C(z)$ —that is, two control variables and a single noise variable. Suppose the response model that is fitted is given by

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + cz + b_{12}x_1x_2 + b_{1z}x_1z + b_{2z}x_2z + b_{12z}x_1x_2z$$

Suppose all terms are significant in the analysis.

(a) Write out the form of the mean model.

(b) Write out the form of the variance model.

(c) What assumptions are being made in the construction of the models in (a) and (b)?

**11.12** Suppose a study is conducted with three control variables and two noise variables. A  $2^{5-1}$  fractional factorial is used as the experimental design. Only the main effects and the  $x_1z_1$ ,  $x_2z_1$ , and  $x_2z_2$  interactions are found to be important. The interaction plots are shown in Fig. E11.1. The main effect plots are shown in Fig. E11.2. The purpose of the experiment is to determine condition on the control variables that *minimize mean response*.

(a) Give approximate conditions on  $x_1$ ,  $x_2$ , and  $x_3$  that result in minimum mean response.

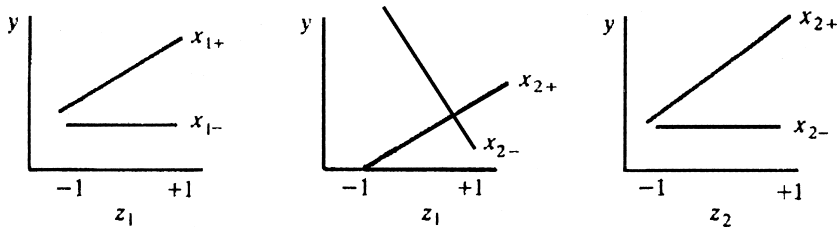


Figure E11.1 Interaction plots for Exercises 11.12 and 11.13.

- (b) Give approximate conditions on  $x_1$ ,  $x_2$ , and  $x_3$  that give approximate minimum process variance.
- (c) Is there a trade-off between optimum conditions? Explain.

11.13 Consider Exercise 11.12. Write out the form of the variance model.

11.14 A  $2^4$  factorial design is used to study a nitride etch process in a single-wafer plasma etcher. The response of interest is the etch rate, and the design factors are

- A, or  $x_1$ : Spacing (gap) between anode and cathode (control factor)
- B, or  $x_2$ : Pressure in the reactor chamber (control factor)
- C, or  $z_1$ : Flow rate of the reactant gas (noise factor)
- D, or  $z_2$ : Power applied to the cathode (noise factor)

The factor levels are given by

|          | $x_1$ | $x_2$ | $z_1$ | $z_2$ |
|----------|-------|-------|-------|-------|
| Low(-1)  | 0.80  | 450   | 125   | 275   |
| High(+1) | 1.20  | 550   | 200   | 325   |

The design and observed results are given in Table E11.1.

- (a) Fit a response model to the etch rate data.
- (b) Find appropriate models for the mean and variance.
- (c) Suppose that it is important to keep the standard deviation of the etch rate below  $150^\circ$ . What operating conditions would you recommend?

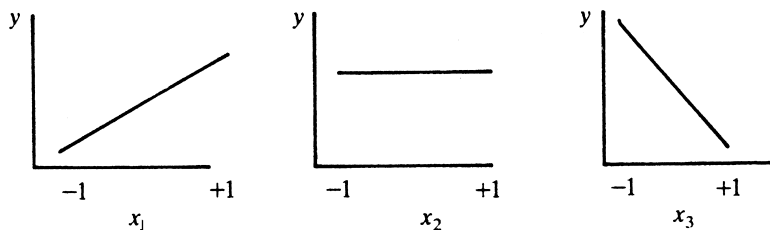


Figure E11.2 Main effect plots for Exercises 11.12 and 11.13.



TABLE E11.1 Data for Exercise 11.14

| $A = x_1$ | $B = x_2$ | $C = z_1$ | $D = z_2$ | $y$ (etch rate, Å/min) |
|-----------|-----------|-----------|-----------|------------------------|
| -1        | -1        | -1        | -1        | 550                    |
| 1         | -1        | -1        | -1        | 669                    |
| -1        | 1         | -1        | -1        | 604                    |
| 1         | 1         | -1        | -1        | 650                    |
| -1        | -1        | 1         | -1        | 633                    |
| 1         | -1        | 1         | -1        | 642                    |
| -1        | 1         | 1         | -1        | 601                    |
| 1         | 1         | 1         | -1        | 635                    |
| -1        | -1        | -1        | 1         | 1037                   |
| 1         | -1        | -1        | 1         | 749                    |
| -1        | 1         | -1        | 1         | 1052                   |
| 1         | 1         | -1        | 1         | 868                    |
| -1        | -1        | 1         | 1         | 1075                   |
| 1         | -1        | 1         | 1         | 860                    |
| -1        | 1         | 1         | 1         | 1063                   |
| 1         | 1         | 1         | 1         | 729                    |

- 11.15** Consider the etch rate data in Exercise 11.14. Show that the condition for minimum process variance is

$$x_1 = \frac{153.0625}{76.8125} = 1.99$$

To minimize the variance, the constraint induced by the experimental region forces the recommendation that one should operate at  $x_1 = 1.0$ .

- (a) Find a confidence region on the location (in  $x_1$ ) of minimum process variance. Does the region cover  $x_1 = 1.0$ ?
- (b) Comment on your findings in (a).
- 11.16** Consider Exercise 3.9. Investigate the data set for dispersion effects by fitting a log variance model. Comment.
- 11.17** Suppose a crossed array of the type

$$2_{\text{III}}^{6-3} \times 2_{\text{III}}^{3-1}$$

is being proposed. Suppose you wish to use two levels on each variable but you wish to have access to more two-factor interactions among control variables. In addition, you wish to fit a response model involving control and noise variables. Suggest an alternative design that better suits the experimenter's priorities.

- 11.18** Consider a situation with three control variables and three noise variables. In order to construct an appropriate response model, the design should contain linear main effects, two-factor interactions among the control variables, and two-factor interactions between control and noise variables.
- (a) What is the minimum number of design points required?
- (b) Give a 22-point design that is efficient for fitting this response function.
- (c) Give an efficient 25-point design.

- 11.19** There are situations in which two-factor interactions involving the noise variables are important, though response models involving these interactions are not emphasized in this text. Consider the fitted model

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + c_1z_1 + c_2z_2 + c_{12}z_1z_2$$

Assume that  $\sigma_{z_1}$  and  $\sigma_{z_2}$  are both 1.0. In addition, assume that  $z_1$  and  $z_2$  are uncorrelated in the process.

- (a) Using the notation in the above fitted model, What is the estimate of the mean model?
- (b) What is the estimate of the variance model?
- (c) Are there any additional assumptions made in (b)?
- 11.20** Reconsider Exercise 11.18. Suppose it is known that factor  $A$  (control factor) is very likely to interact with noise variables  $D$  and  $E$ . All control variables  $A$ ,  $B$ , and  $C$  are likely to be involved in two-factor interactions. Design a combined array that allows for an appropriate robust parameter design. Use 16 observations with a regular fraction of a  $2^6$ . Give appropriate defining relations.
- 11.21** Consider the data in Example 6.15 with the use of the conversion response. Suppose, in fact, that time and catalyst are control variables and temperature is a noise variable. Use a response model to produce the following:
- (a) A process mean model
- (b) A process variance model
- What assumptions did you make in developing the above models?
- 11.22** As we indicated in Section 11.4.4, the CCD can be very useful. Suppose there are four control variables and four noise Variables. Give a CCD that is a combined array and allows for estimation of a second-order model in the control variables, all control  $\times$  noise interactions and noise main effect terms. This design should contain 64 factorial points and eight axial points, plus center runs. Use  $\alpha = 1.0$ .
- 11.23** Construct a  $D$ -optimal design to accommodate the scenarios described in Exercise 11.22. Use 45 total design points. Repeat using an  $I$ -optimal design. Compare the two designs.
- 11.24** Consider a situation with the three control and two noise variables to accommodate a model that is second-order in the control variables and involves linear noise variable terms as well as all control  $\times$  noise interactions.
- (a) Give an appropriate CCD ( $\alpha = 1.0$ ) containing 22 points plus center runs.
- (b) Give an appropriate  $D$ -optimal design.
- (c) Give an appropriate  $I$ -optimal design.
- 11.25** Consider Example 11.1.
- (a) Fit an appropriate response model.
- (b) Use the response model to generate a model for the process mean and a model for the process variance.
- (c) Compare your results with those in Example 11.1 and with the direct variance modeling approach in Example 11.7

**TABLE E11.2** Data for Exercise 11.26

| Control Variables |       |       | Noise Variables for the Following Values of $z_1$ and $z_2$ |                      |                      |                      | Response Average | Response Standard Deviation | Type I Signal-to-Noise Ratio |
|-------------------|-------|-------|---|----------------------|----------------------|----------------------|------------------|-----------------------------|------------------------------|
|                   |       |       | $z_1 = -1, z_2 = -1$  | $z_1 = +1, z_2 = -1$ | $z_1 = -1, z_2 = +1$ | $z_1 = +1, z_2 = +1$ |                  |                             |                              |
| $x_1$             | $x_2$ | $x_3$ |   |                      |                      |                      |                  |                             |                              |
| -1                | -1    | -1    | 118.9   | 65.7                 | 95.3                 | 92.4                 | 93.08            | 21.77                       | 29.06                        |
| +1                | -1    | -1    | 153.7   | 229.4                | 119.9                | 251.5                | 188.63           | 62.07                       | 31.35                        |
| -1                | +1    | -1    | 196.7   | 170.9                | 234.2                | 166.6                | 192.10           | 31.06                       | 36.44                        |
| +1                | +1    | -1    | 211.1   | 245.7                | 241.0                | 252.6                | 237.60           | 18.30                       | 51.28                        |
| -1                | -1    | +1    | 145.2   | 132.2                | 167.1                | 137.9                | 145.60           | 15.29                       | 45.07                        |
| +1                | -1    | +1    | 125.3   | 201.6                | 185.5                | 267.3                | 194.93           | 58.36                       | 24.11                        |
| -1                | +1    | +1    | 283.0   | 251.1                | 263.4                | 190.4                | 246.98           | 39.94                       | 36.44                        |
| +1                | +1    | +1    | 184.2   | 279.5                | 247.2                | 259.2                | 242.53           | 41.11                       | 35.50                        |

- 11.26** Montgomery (1999) describes an experiment performed in a semiconductor factory to determine how five factors influence the transistor gain for a particular device. It is desirable to hold the gain as close as possible to a target value of 200 (the specifications are actually  $200 \pm 20$ ). Three of the variables are relatively easy to control:  $x_1$  (implant dose),  $x_2$  (drive-in time), and  $x_3$  (vacuum level). Two variables are difficult to control in routine manufacturing and are considered as noise factors:  $z_1$  (oxide thickness) and  $z_2$  (temperature). The crossed array design they used is shown in Table E11.2. Conduct a Taguchi-type analysis to determine the operating conditions.
- 11.27** Reconsider the experiment described in Example 11.26. Treat the outer array runs as replicates, and fit models for the mean and variance using the control variables. Find the optimum operating conditions for the process. How do they compare with the results from the Taguchi-type analysis in Exercise 11.26?
- 11.28** Reconsider the experiment described in Exercise 11.26. Fit a response model to the data involving both the control and noise variables. Derive the mean and variance models, and find optimum operating conditions. Compare your results with those obtained in Exercises 11.26 and 11.27.
- 11.29** The experiment in Exercise 11.26 is a  $2^3 \times 2^2$  crossed array. Viewed as a combined array, the design is a  $2^5$ . Using only the runs that correspond to a  $2^{5-1}$ , rework Exercise 11.28. Does this tell you anything useful about combined array designs?
- 11.30** Consider the experiment in Exercise 6.15. Suppose that temperature is a noise variable ( $\sigma_z^2 = 1$  in coded units). Fit response models for both responses. Is there a robust design problem with respect to both responses? Find a set of conditions that maximize conversion with activity between 55 and 60 and that minimize the variability transmitted from temperature.
- 11.31** An experiment has been run in a process that applies a coating material to a wafer. Each run in the experiment produced a wafer, and the coating thickness was measured several times at different locations on the wafer. Then the mean  $y_1$

TABLE E11.3 Data for Exercise 11.31

| Run | Speed | Pressure | Distance | Mean $y_1$ | Std. Dev. $y_2$ |
|-----|-------|----------|----------|------------|-----------------|
| 1   | -1    | -1       | -1       | 24.0       | 12.5            |
| 2   | 0     | -1       | -1       | 120.3      | 8.4             |
| 3   | 1     | -1       | -1       | 213.7      | 42.8            |
| 4   | -1    | 0        | -1       | 86.0       | 3.5             |
| 5   | 0     | 0        | -1       | 136.6      | 80.4            |
| 6   | 1     | 0        | -1       | 340.7      | 16.2            |
| 7   | -1    | 1        | -1       | 112.3      | 27.6            |
| 8   | 0     | 1        | -1       | 256.3      | 4.6             |
| 9   | 1     | 1        | -1       | 271.7      | 23.6            |
| 10  | -1    | -1       | 0        | 81.0       | 0.0             |
| 11  | 0     | -1       | 0        | 101.7      | 17.7            |
| 12  | 1     | -1       | 0        | 357.0      | 32.9            |
| 13  | -1    | 0        | 0        | 171.3      | 15.0            |
| 14  | 0     | 0        | 0        | 372.0      | 0.0             |
| 15  | 1     | 0        | 0        | 501.7      | 92.5            |
| 16  | -1    | 1        | 0        | 264.0      | 63.5            |
| 17  | 0     | 1        | 0        | 427.0      | 88.6            |
| 18  | 1     | 1        | 0        | 730.7      | 21.1            |
| 19  | -1    | -1       | 1        | 220.7      | 133.8           |
| 20  | 0     | -1       | 1        | 239.7      | 23.5            |
| 21  | 1     | -1       | 1        | 422.0      | 18.5            |
| 22  | -1    | 0        | 1        | 199.0      | 29.4            |
| 23  | 0     | 0        | 1        | 485.3      | 44.7            |
| 24  | 1     | 0        | 1        | 673.7      | 158.2           |
| 25  | -1    | 1        | 1        | 176.7      | 55.5            |
| 26  | 0     | 1        | 1        | 501.0      | 138.9           |
| 27  | 1     | 1        | 1        | 1010.0     | 142.4           |

and standard deviation  $y_2$  of the thickness measurement were obtained. The data [adapted from Box and Draper (1987)] are shown in Table E11.3.

- What type of design did the experimenters use? Is this a good choice of design for fitting a quadratic model?
- Build models of both responses.
- Find a set of optimum conditions that result in a mean as large as possible with the standard deviation less than 60.

**11.32** In Example 11.3 we found that one of the process variables ( $B$  = pressure) was not important. Dropping this variable produces two replicates of a  $2^3$  design. The data are shown below:

| $C$ | $D$ | $A(+)$ | $A(-)$   | $\bar{y}$ | $s^2$   |
|-----|-----|--------|----------|-----------|---------|
| -   | -   | 45, 48 | 71, 65   | 57.75     | 121.19  |
| +   | -   | 68, 80 | 60, 65   | 68.25     | 72.25   |
| -   | +   | 43, 45 | 100, 104 | 73.00     | 1124.67 |
| +   | +   | 75, 70 | 86, 96   | 81.75     | 134.92  |

Assume that  $C$  and  $D$  are controllable factors and that  $A$  is a noise variable.

- (a) Fit a model to the mean response.
- (b) Fit a log model to the variance.
- (c) Find operating conditions that result in the mean filtration rate response exceeding 75 with minimum variance.
- (d) Compare your results with those from Example 11.3, which used the response model approach. How similar are the two answers?

**11.33** Consider the response model

$$\hat{y} = 24 + 16x_1 + 12x_2 + 4z - 2x_1z + 3x_2z - 2x_1x_2z$$

The design used to fit the model was a  $2^3$  factorial in the usual coded units. All terms in the model are significant.

- (a) Write out the mean model
- (b) Write out the variance model.
- (c) What value of  $z$  will minimize the transmitted variance?

**11.34** Consider the response model

$$\hat{y} = 10 + 8x_1 + 6x_2 + 3z + 2x_1z + 1x_2z$$

The design used to fit the model was a  $2^3$  factorial in the usual coded units. All terms in the model are significant. The experimenter wants to find levels of the design variables and the noise variable that maximize the response and make the transmitted variability from the noise variable  $z$  as small as possible. What operating conditions do you recommend?

**11.35** Consider the response model

$$\hat{y} = 15 + 6x_1 + 10x_2 + 6z - 2x_1z + 5x_2z - 2x_1x_2z + 3x_1^2$$

The design used to fit the model was a central composite design in the usual coded units. The original model was a full quadratic and it was reduced leaving only the terms in the model above.

- (a) Write out the model for the mean.
- (b) Write out the model for the transmitted variance.
- (c) Suppose that the objective is to maximize the response while making the transmitted variance as small as possible. What set of operating conditions do you recommend?

**11.36** We have assumed that noise variables are controllable for purposes of an experiment. What strategy would you recommend if the noise variables were not controllable but their values could be observed while the experiment in terms of the controllable process variables was being conducted?

**Short Answer Questions**

- 11.37** Signal-to-noise factors generally confound location and dispersion effects. **True False**
- 11.38** A crossed array design allows estimation of the control-by-noise factor interactions. **True False**
- 11.39** If a  $3^{3-1}$  design is used as the inner array and a  $2^3$  is used as the outer array, all interactions can be estimated. **True False**
- 11.40** Suppose that the response model is  $\hat{y} = 6 + 4x_1 + 3x_2 + 3z + 2x_1z + 2x_2z$ . The value of the controllable factor  $x_1$  that minimizes the transmitted variance is  
(a)  $-1$   
(b)  $+1$   
(c)  $-0.5$   
(d) None of the above
- 11.41** Consider the response model in Exercise 11.40. The value of the controllable factor  $x_2$  that minimizes the transmitted variance is  
(a)  $-1$   
(b)  $+1$   
(c)  $-0.5$   
(d) None of the above
- 11.42** A design that allows estimation of control-by-control factor and control-by-noise factor interactions but may not allow estimation of noise-by-noise factor interactions is called a mixed resolution design. **True False**
- 11.43** Residuals can sometimes be used to estimate dispersion effects in unreplicated and/or fractionated designs. **True False**
- 11.44** A log transformation is frequently useful in directly modeling the variance as a response. **True False**
- 11.45** Computer-generated optimal designs can be a useful approach for designing combined array experiments with both control and noise variables. **True False**

---

# 12

---

## EXPERIMENTS WITH MIXTURES

### 12.1 INTRODUCTION

In the response surface examples discussed previously, the levels chosen for any factor in the experimental design are independent of the levels chosen for the other factors. For example, suppose that there are three variables—stirring rate, reaction time, and temperature—that affect the yield of a chemical process. The levels of temperature may be chosen independently of the levels of reaction time and stirring rate, and consequently we may think of the region of experimentation as either a cube or a sphere. The experimental design will consist of an appropriate set of points over this cuboidal or spherical region (such as a factorial design or a central composite design).

A **mixture experiment** is a special type of response surface experiment in which the factors are the ingredients or components of a mixture, and the response is a function of the proportions of each ingredient. These proportional amounts of each ingredient are typically measured by weight, by volume, by mole ratio, and so forth.

For example, consider chemical etching of semiconductor wafers. The process consists of placing the wafers in a solution composed of several acids—say nitric acid, hydrochloric acid, and phosphoric acid—and observing the etch rate. In such a process, we wish to determine if a combination of these acids is more effective in etching the wafers than any of the three acids by themselves. Because the volume of the etching chamber is fixed, an experiment might consist of testing various combinations of these acids, where all mixes or blends have the same volume. If  $x_1$ ,  $x_2$ , and  $x_3$  represent the proportions by volume of the three acids, then some of the blends that might be of interest are as follows:

$$\text{Blend 1: } x_1 = 0.3, \quad x_2 = 0.3, \quad x_3 = 0.4$$

$$\text{Blend 2: } x_1 = 0.2, \quad x_2 = 0.5, \quad x_3 = 0.3$$

Blend 3:  $x_1 = 0.5, \quad x_2 = 0.5, \quad x_3 = 0.0$

Blend 4:  $x_1 = 1.0, \quad x_2 = 0.0, \quad x_3 = 0.0$

Blends 1 and 2 are examples of **complete** mixtures; that is, they are made up of all three of the acids. Blend 3 is a **binary** blend, consisting of two of the three acids, and blend 4 is a mixture that is made up of 100% by volume of only one of the three acids. Each of these blends is called a **mixture**. Notice that in this example  $x_1 + x_2 + x_3 = 1$ ; and because of this constraint, the levels of the factors cannot be chosen independently. For instance, in blend 1 above, as soon as we indicate that acid 1 makes up 30% of the mixture by volume ( $x_1 = 0.3$ ) and acid 2 makes up 30% of the mixture by volume ( $x_2 = 0.3$ ), it is immediately obvious that acid 3 must make up 40% of the mixture by volume (that is,  $x_3 = 0.4$  because  $x_1 = x_2 = 0.3$  and  $x_1 + x_2 + x_3 = 1.0$ ).

In general, suppose that the mixture consists of  $q$  ingredients or components, and let  $x_i$  represent the proportion of the  $i$ th ingredient in the mixture. Then in light of the above discussion we must require that

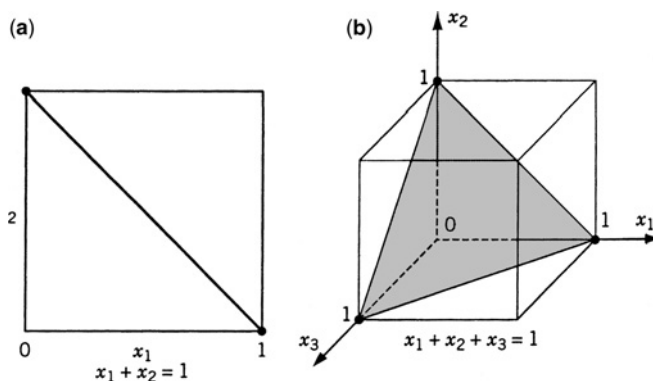
$$x_i \geq 0, \quad i = 1, 2, \dots, q \quad (12.1)$$

and

$$\sum_{i=1}^q x_i = x_1 + x_2 + \dots + x_q = 1 \quad (12.2)$$

The constraint in Equation 12.2 makes the levels of the factors  $x_i$  dependent and this makes mixture experiments different from the usual response surface experiments we have discussed previously.

The constraints in Equations 12.1 and 12.2 are shown graphically in Fig. 12.1 for  $q = 2$  and  $q = 3$  components. For two components, the feasible factor space for the mixture experiment includes all values of the two components for which  $x_1 + x_2 = 1$ , which is the line segment connecting the two points  $x_1 = 0, x_2 = 1$ , and  $x_1 = 1, x_2 = 0$  shown in Fig. 12.1a. With three components, the feasible space for the mixture experiment is the triangle in Fig. 12.1b that has vertices corresponding to the three **pure blends**—that is,



**Figure 12.1** Constrained factor space for mixtures with (a)  $q = 2$  components and (b)  $q = 3$  components.



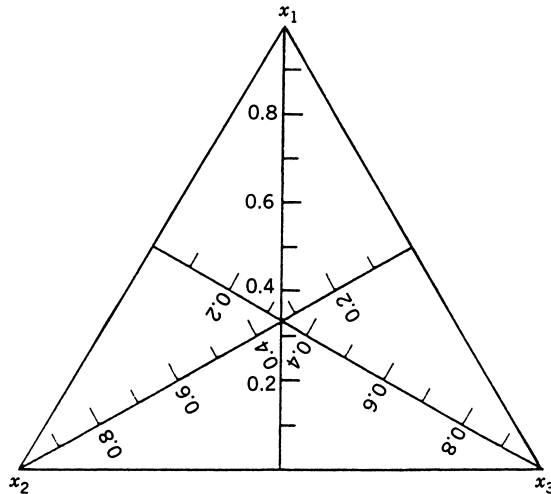


Figure 12.2 Simplex coordinate system for three components.

mixtures that are made up of 100% of a single ingredient—and edges that are binary blends. In general, the experimental region for a mixture problem with  $q$  components is a **simplex**, which is a regularly sided figure with  $q$  vertices in  $q - 1$  dimensions.

The coordinate system for mixture proportions is a **simplex coordinate system**. For example, with  $q = 3$  components, the experimental region is shown in Fig. 12.2. Each of the three vertices in the equilateral triangle corresponds to a pure blend, and each of the three sides of the triangle represents a binary blend mixture that has none of one of the three components (the component labeled on the opposite vertex). The nine grid lines in each direction mark off 10% increments in the respective components. Interior points in the triangle represent mixtures in which all three ingredients are present at nonzero proportionate amounts. The centroid of the triangle corresponds to the mixture with equal proportions  $x_1 = \frac{1}{3}$ ,  $x_2 = \frac{1}{3}$ ,  $x_3 = \frac{1}{3}$  of all ingredients.

Applications of mixture experiments are found in many areas. A common application is **product formulation**, in which a product is formed by mixing several ingredients together. Examples include (a) formulation of gasoline by combining several refinery products and other types of chemicals such as anticlers, (b) formulation of soaps, shampoos, and detergents in the consumer products industry, and (c) formulation of products such as cake mixes and beverages in the food industry. In addition to product formulation, mixture experiments arise frequently in process engineering and development where some property of a final product depends on a mixture of several ingredients interacting with the product at a particular process stage. An example is the wafer etching process described above. Other examples include gas or plasma etching in the semiconductor industry, many joining processes such as soldering, and chemical milling or machining.

In this chapter we discuss experimental design, data analysis, and model-building techniques for the original mixture experiment defined by Equations 12.1 and 12.2. In this problem, the entire simplex region is investigated, and the only design variables are the mixture components  $x_1, x_2, \dots, x_q$ .

There are many variations of the original mixture problem. One of these is the addition of upper and lower bounds on some of the component proportions. These bounds occur because

the problem may require that at least a certain minimum proportion of an ingredient be present in the blend, or that an ingredient cannot exceed a specified maximum proportion. In some mixture problems there are **process variables**  $z_1, z_2, \dots, z_p$  in addition to the mixture ingredients. For example, in the wafer etching example above, the mixture ingredients  $x_1, x_2,$  and  $x_3$  are the proportions of the three acids, and two process variables could be the temperature ( $z_1$ ) of the etching solution and the agitation rate ( $z_2$ ) of the wafers. Including process variables in mixture experiments usually involves constructing designs that utilize different levels of the process variables via a factorial design in combination with the mixture variables. Finally, in some mixture problems the experimenter may be interested in not only the formulation of the mixture but the amount of the mixture that is applied to the experimental units. For example, the opacity of a coating depends on the ingredients from which the coating is formulated, but it also depends upon the thickness of the coating surface when it is applied. Some of these variations of the original mixture experiment will be discussed in Chapter 13.

## 12.2 SIMPLEX DESIGNS AND CANONICAL MIXTURE POLYNOMIALS

The primary differences between a standard response surface experiment and a mixture experiment are that (1) special types of designs must be used and (2) the form of the mixture polynomial model is slightly different from the standard polynomial models used in response surface methodology (RSM). In this section we introduce designs that allow an appropriate response surface model to be fitted over the entire mixture space. Because the mixture space is a simplex, all design points must be at the vertices, on the edges or faces, or in the interior of a simplex.

Much of the work in this area was originated by Scheffé (1958, 1959, 1965). Cornell (2002) is an excellent and very complete reference on the subject.

### 12.2.1 Simplex Lattice Designs

A simplex lattice is just a uniformly spaced set of points on a simplex. A  $\{q, m\}$  simplex lattice design for  $q$  components consists of points defined by the following coordinate settings: The proportions taken on by each component are the  $m + 1$  equally spaced values from 0 to 1,

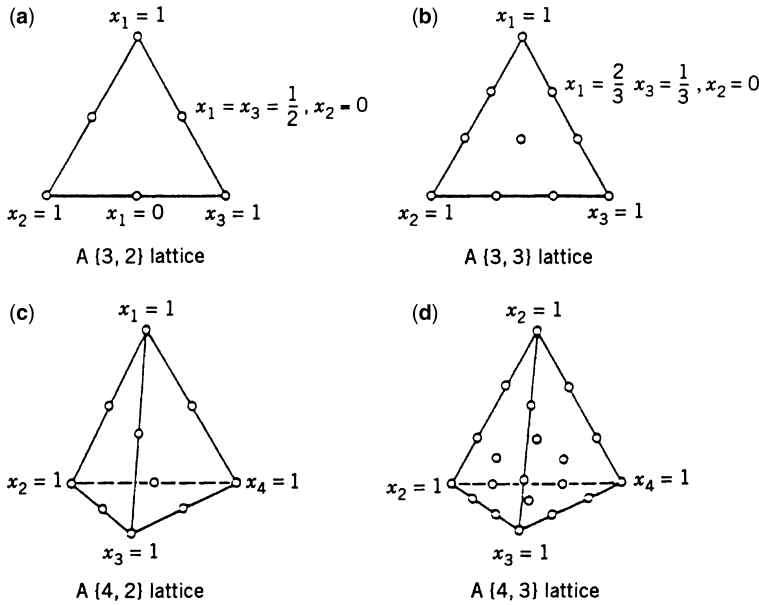
$$x_i = 0, \frac{1}{m}, \frac{2}{m}, \dots, 1, \quad i = 1, 2, \dots, q \quad (12.3)$$

and all possible combinations (mixtures) of the proportions from this equation are used. As an example, let  $q = 3$  and  $m = 2$ ; consequently,

$$x_i = 0, \frac{1}{2}, 1, \quad i = 1, 2, 3$$

and the simplex lattice consists of the following six points:

$$(x_1, x_2, x_3) = (1, 0, 0), (0, 1, 0), (0, 0, 1), \left(\frac{1}{2}, \frac{1}{2}, 0\right), \left(\frac{1}{2}, 0, \frac{1}{2}\right), \left(0, \frac{1}{2}, \frac{1}{2}\right)$$



**Figure 12.3** Simplex lattice designs for  $q = 3$  and  $q = 4$  components.

This is a  $\{3, 2\}$  simplex lattice design, and it is shown in Fig. 12.3a. The three vertices  $(1, 0, 0)$ ,  $(0, 1, 0)$ , and  $(0, 0, 1)$  are pure blends, and the points  $(\frac{1}{2}, \frac{1}{2}, 0)$ ,  $(\frac{1}{2}, 0, \frac{1}{2})$ , and  $(0, \frac{1}{2}, \frac{1}{2})$  are binary blends or two-component mixtures located at the midpoints of the three edges of the triangle. These binary blends are made up of equal parts of two of the three ingredients. Figure 12.3 also shows the  $\{3, 3\}$ , the  $\{4, 2\}$ , and the  $\{4, 3\}$  simplex lattice designs. As we will subsequently see, the notation  $\{q, m\}$  implies a simplex lattice design in  $q$  components that will support a mixture polynomial of degree  $m$ . In general, the number of points in a  $\{q, m\}$  simplex lattice design is

$$N = \frac{(q + m - 1)!}{m!(q - 1)!} \tag{12.4}$$

Now consider the form of the polynomial model that we might fit to data from a mixture experiment. A first-order model is

$$E(y) = \beta_0 + \sum_{i=1}^q \beta_i x_i \tag{12.5}$$

However, because of the restriction  $x_1 + x_2 + \dots + x_q = 1$  we know that the parameters  $\beta_0, \beta_1, \dots, \beta_q$  are not unique. We could, of course, make the substitution

$$x_q = 1 - \sum_{i=1}^{q-1} x_i$$

in Equation 12.5, removing the dependence among the  $x_i$ -terms and producing unique estimates of the parameters  $\beta_0, \beta_1, \dots, \beta_{q-1}$ . Although this will work mathematically, it is

not the preferred approach, because it obscures the effect of the  $q$ th component in that the term  $\beta_q x_q$  is not included in the equation. Thus, in a sense, we are sacrificing information on component  $q$ . This model is known as a **slack variable** model, described further in Section 12.4.

An alternative approach is to multiply some of the terms in the original response surface polynomial by the identity  $x_1 + x_2 + \dots + x_q = 1$  and then simplify. For example, consider Equation 12.5, and multiply the term  $\beta_0$  by  $x_1 + x_2 + \dots + x_q = 1$ , yielding

$$\begin{aligned}
 E(y) &= \beta_0(x_1 + x_2 + \dots + x_q) + \sum_{i=1}^q \beta_i x_i \\
 &= \sum_{i=1}^q \beta_i^* x_i
 \end{aligned}
 \tag{12.6}$$

where  $\beta_i^* = \beta_0 + \beta_i$ . This is called the **canonical** or **Scheffé form** of the first-order mixture model. In general, the canonical or Scheffé forms of the mixture models (with the asterisks removed from the parameters) are as follows:

*Linear:*

$$E(y) = \sum_{i=1}^q \beta_i x_i
 \tag{12.7}$$

*Quadratic:*

$$E(y) = \sum_{i=1}^q \beta_i x_i + \sum_{i<j=2}^q \sum \beta_{ij} x_i x_j
 \tag{12.8}$$

*Full Cubic:*

$$\begin{aligned}
 E(y) &= \sum_{i=1}^q \beta_i x_i + \sum_{i<j=2}^q \sum \beta_{ij} x_i x_j + \sum_{i<j=2}^q \sum \delta_{ij} x_i x_j (x_i - x_j) \\
 &\quad + \sum_{i<j<k=3}^q \sum \sum \beta_{ijk} x_i x_j x_k
 \end{aligned}
 \tag{12.9}$$

*Special Cubic:*

$$E(y) = \sum_{i=1}^q \beta_i x_i + \sum_{i<j=2}^q \sum \beta_{ij} x_i x_j + \sum_{i<j<k=3}^q \sum \sum \beta_{ijk} x_i x_j x_k
 \tag{12.10}$$

The terms in the canonical mixture polynomials have simple interpretations. Geometrically, in Equations 12.7 through 12.10, the parameter  $\beta_i$  represents the expected response for the pure mixture  $x_i = 1, x_j = 0, j \neq i$ , and is the height of the mixture surface at the vertex  $x_i = 1$ . The portion of each polynomial given by

$$\sum_{i=1}^q \beta_i x_i$$

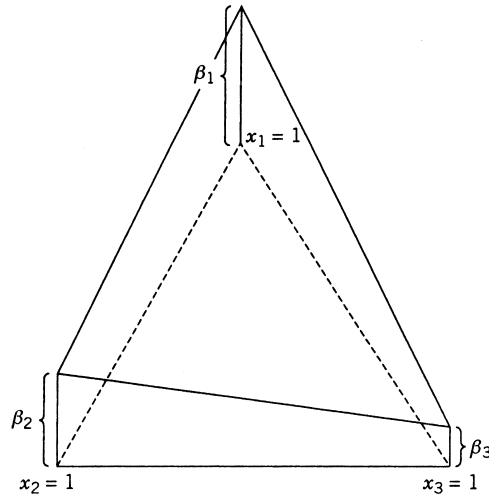


Figure 12.4 Linear mixture model in three components with  $\beta_1 > \beta_2 > \beta_3$ .

is called the **linear blending** portion. When blending is strictly additive, then Equation 12.7 is an appropriate model. Figure 12.4 presents the case where  $q = 3$  and  $\beta_1 > \beta_2 > \beta_3$ .

Quadratic blending is illustrated in Fig. 12.5a. Notice that the quadratic term  $\beta_{12} x_1 x_2$  represents the excess response from the quadratic model  $E(y) = \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$  over the linear model. This is often called the **synergism** (or **antagonism**) **due to nonlinear blending**. For example, if large positive values of  $y$  are desired and if  $\beta_{12}$  is positive, synergistic blending is occurring, whereas if  $\beta_{12}$  is negative, antagonistic blending is occurring. In the cubic model, terms such as  $\delta_{12} x_1 x_2 (x_1 - x_2)$  enable one to model both synergistic and antagonistic blending along the  $x_1 - x_2$  edge; see Fig. 12.5b. A cubic term such as  $\beta_{123} x_1 x_2 x_3$  accounts for ternary blending among the three components in the interior of the simplex.

It is also helpful to consider the way in which the individual terms in a mixture model contribute to the shape of the response surface. A linear term such as  $\beta_1 x_1$  only contributes to the model when  $x_1 > 0$ ; and the maximum contribution occurs at  $x_1 = 1$ , in which case the maximum effect contributed by  $x_1$  is  $\beta_1$ . The quadratic term  $\beta_{12} x_1 x_2$  contributes to the model at every point in the simplex where  $x_1 > 0$  and  $x_2 > 0$ . The maximum contribution occurs at the edge joining the vertices  $x_1$  and  $x_2$  and is at the point  $x_1 = x_2 = \frac{1}{2}$ .

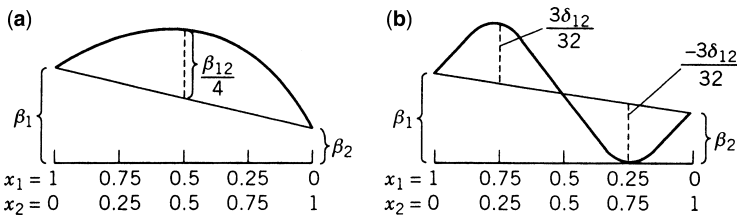


Figure 12.5 Nonlinear blending. (a) Quadratic blending with  $\beta_{ij} > 0$ . (b) Cubic blending with  $\delta_{12} > 0$ .

The maximum contribution to the model from this term is  $\beta_{12}/4$ . A cubic term such as  $\beta_{123}x_1x_2x_3$  contributes to the model at every point for which  $x_1 > 0$ ,  $x_2 > 0$ , and  $x_3 > 0$  (in the interior of the simplex), and the maximum contribution is of magnitude  $\beta_{123}/27$  at the point  $x_1 = x_2 = x_3 = \frac{1}{3}$ .

**Example 12.1 A Three-Component Mixture** Cornell (2002) describes a mixture experiment in which three components—polyethylene ( $x_1$ ), polystyrene ( $x_2$ ), and polypropylene ( $x_3$ )—were blended to form fiber that will be spun into yarn for draperies. The product developers were interested in only the pure and binary blends of these three materials. The response variable of interest is yarn elongation in kilograms of force applied. A  $\{3, 2\}$  simplex lattice design is used to study the product. The design and the observed responses are shown in Table 12.1. Notice that replicate observations are run, with two replicates at each of the pure blends and three replicates at each of the binary blends. The error standard deviation can be estimated from these replicate observations as  $\hat{\sigma} = 0.85$ .

Both Design-Expert and JMP have excellent capability for design and analysis of mixture experiments. Table 12.2 shows output from the Design-Expert computer program, which was used to analyze the yarn elongation data. This program fits both the linear and quadratic mixture model to the data, and the upper panel of Table 12.2 shows the sequential  $F$ -tests for the linear blending terms and the quadratic terms. The  $F$ -test for the linear terms is significant, and the  $F$ -test for the contribution of the quadratic terms over the linear terms is also significant, indicating that a quadratic mixture model should be used. The middle panel of Table 12.2 shows a lack-of-fit test for the linear mixture model. The value of  $F = 32.32$  is very large, indicating that a linear mixture model is inadequate. Finally, the bottom panel of Table 12.2 presents several informative summary statistics for the linear and quadratic models. These statistics indicate that the quadratic model is superior to the linear model.

The  $F$ -test for the linear terms in the upper panel of Table 12.2 warrants some discussion. This is a test of the hypothesis that there is no linear blending occurring in the mixture. If there is no linear blending, then the response surface is a level plane above the simplex region; that is,  $\beta_1 = \beta_2 = \beta_3 = \beta$ . Expressed formally, the hypotheses to be tested are

$$\begin{aligned} H_0: & \quad \beta_1 = \beta_2 = \beta_3 = \beta \\ H_1: & \quad \text{At least one equality is false} \end{aligned}$$

Because from Table 12.2 the test statistic is  $F_{\text{linear}} = 4.477$  with a  $P$ -value of  $P = 0.0353$ , we would reject  $H_0$  and conclude that there is linear blending in the system, in other words it makes a difference which fibers are used in the yarn.

**TABLE 12.1 The  $\{3, 2\}$  Simplex Lattice Design for the Yarn Elongation Problem**

| Design Point | Component Proportions |               |               | Observed Elongation Values | Average Elongation Value $\bar{y}$ |
|--------------|-----------------------|---------------|---------------|----------------------------|------------------------------------|
|              | $x_1$                 | $x_2$         | $x_3$         |                            |                                    |
| 1            | 1                     | 0             | 0             | 11.0, 12.4                 | 11.7                               |
| 2            | $\frac{1}{2}$         | $\frac{1}{2}$ | 0             | 15.0, 14.8, 16.1           | 15.3                               |
| 3            | 0                     | 1             | 0             | 8.8, 10.0                  | 9.4                                |
| 4            | 0                     | $\frac{1}{2}$ | $\frac{1}{2}$ | 10.0, 9.7, 11.8            | 10.5                               |
| 5            | 0                     | 0             | 1             | 16.8, 16.0                 | 16.4                               |
| 6            | $\frac{1}{2}$         | 0             | $\frac{1}{2}$ | 17.7, 16.4, 16.6           | 16.9                               |

**TABLE 12.2 Design-Expert Output for the Yarn Elongation Data**

| <i>Sequential Model Sum of Squares</i>       |                 |                             |                         |           |                    |        |
|--|-----------------|-----------------------------|-------------------------|-----------|--------------------|--------|
| Source                                       | Sum of Squares  | Degrees of Freedom          | Mean Square             | F-Value   | Probability > F    |        |
| Mean   | 2750.0          | 1                           | 2750.0                  |           |                    |        |
| Linear                                       | 57.6            | 2                           | 28.8                    | 4.477     | 0.0353             |        |
| Quadratic                                    | 70.7            | 3                           | 23.6                    | 32.32     | 0.0001             |        |
| Residual                                     | 6.6             | 9                           | 0.7                     |           |                    |        |
| Total  | 2884.9          | 15                          |                         |           |                    |        |
| <i>Lack-of-Fit Tests</i>                     |                 |                             |                         |           |                    |        |
| Model  | Sum of Squares  | Degrees of Freedom          | Mean Square             | F-Value   | Probability > F    |        |
| Linear                                       | 70.7            | 3                           | 23.6                    | 32.32     | 0.0001             |        |
| Quadratic                                    | 0.0             | 0                           |                         |           |                    |        |
| Pure Error                                   | 6.6             | 9                           | 0.7                     |           |                    |        |
| <i>ANOVA Summary Statistics of Model Fit</i> |                 |                             |                         |           |                    |        |
| Source                                       | Unaliated Terms | Residual Degrees of Freedom | Root Mean Squared Error | R-Squared | Adjusted R-Squared | PRESS  |
| Linear                                       | 3               | 12                          | 2.54                    | 0.4273    | 0.3319             | 120.81 |
| Quadratic                                    | 6               | 9                           | 0.85                    | 0.9514    | 0.9243             | 18.30  |

Table 12.3 summarizes the quadratic model fit. The analysis of variance for the quadratic model, shown in the upper panel of Table 12.3, tests the hypothesis that the response surface is a level plane above the simplex region. Formally, this hypothesis is

$$\begin{aligned}
 H_0: & \quad \beta_1 = \beta_2 = \beta_3 = \beta, \quad \beta_{12} = \beta_{13} = \beta_{23} = 0 \\
 H_1: & \quad \text{At least one equality is false}
 \end{aligned}$$

Because the test statistic is  $F = 35.2$  (with a  $P$ -value of  $P = 0.0001$ ), this hypothesis is rejected. Notice that  $t$ -tests are provided for the quadratic terms on the model because tests of the hypotheses regarding each of these individual terms (i.e.,  $H_0: \beta_{ij} = 0$ ) are meaningful. However, because similar tests on the linear blending terms are not meaningful ( $H_0: \beta_i = 0$  does not test for linear blending from component  $i$ ), these  $t$ -tests are not displayed. Note this quadratic test differ from that in Table 12.2 (upper part with  $F$ -value = 32.32) which just tests if all  $\beta_{ij} = 0$ .

The fitted quadratic mixture model is

$$\hat{y} = 11.7x_1 + 9.4x_2 + 16.4x_3 + 19.0x_1x_2 + 11.4x_1x_3 - 9.6x_2x_3$$

Because  $b_3 > b_1 > b_2$ , we would conclude that component 3 (polypropylene) produces yarn with the highest elongation. Furthermore, because  $b_{12}$  and  $b_{13}$  are positive, blending components 1 and 2 or components 1 and 3 produces higher elongation values than would be expected just by averaging the elongations of the pure blends. This is an example of **synergistic** blending effects. Components 2 and 3 have antagonistic blending effects, because  $b_{23}$  is negative.

Figures 12.6 and 12.7 present a normal probability plot of the studentized residuals from the quadratic model. These plots are satisfactory. In general, we recommend analyzing

TABLE 12.3 Quadratic Mixture Model for the Yarn Elongation Data

| ANOVA for Quadratic Model |                |                    |                    |         |                 |
|---------------------------|----------------|--------------------|--------------------|---------|-----------------|
| Source                    | Sum of Squares | Degrees of Freedom | Mean Square        | F-Value | Probability > F |
| Model                     | 128.3          | 5                  | 25.66              | 35.20   | 0.0001          |
| Residual                  | 6.6            | 9                  | 0.73               |         |                 |
| Corrected total           | 134.9          | 14                 |                    |         |                 |
| Root mean squared error   | 0.85           |                    | R-squared          | 0.9514  |                 |
| Dependent mean            | 13.54          |                    | Adjusted R-squared | 0.9243  |                 |

| <i>t</i> for $H_0$   |                      |                    |                |                 |                          |
|----------------------|----------------------|--------------------|----------------|-----------------|--------------------------|
| Independent Variable | Coefficient Estimate | Degrees of Freedom | Standard Error | Coefficient = 0 | Probability >   <i>t</i> |
| $x_1$                | 11.70                | 1                  | 0.60           | Not applicable  |                          |
| $x_2$                | 9.40                 | 1                  | 0.60           | Not applicable  |                          |
| $x_3$                | 16.40                | 1                  | 0.60           | Not applicable  |                          |
| $x_1x_2$             | 19.00                | 1                  | 2.61           | 7.285           | 0.0001                   |
| $x_1x_3$             | 11.40                | 1                  | 2.61           | 4.371           | 0.0018                   |
| $x_2x_3$             | -9.60                | 1                  | 2.61           | -3.681          | 0.0051                   |

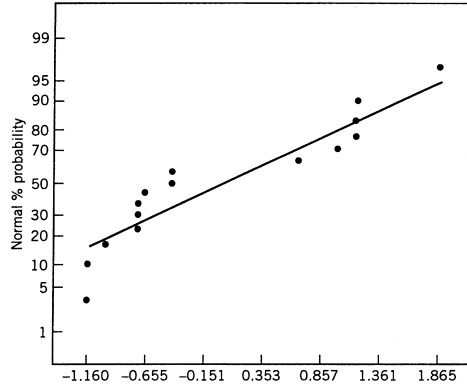
  

| Observation | Actual Value | Predicted Value | Residual | $h_{ii}$ | Studentized Residual | Cook's Distance | R-Student |
|-------------|--------------|-----------------|----------|----------|----------------------|-----------------|-----------|
| 1           | 11.00        | 11.70           | -0.70    | 0.500    | -1.160               | 0.224           | -1.185    |
| 2           | 12.40        | 11.70           | 0.70     | 0.500    | 1.160                | 0.224           | 1.185     |
| 3           | 15.00        | 15.30           | -0.30    | 0.333    | -0.430               | 0.015           | -0.410    |
| 4           | 14.80        | 15.30           | -0.50    | 0.333    | -0.717               | 0.043           | -0.696    |
| 5           | 16.10        | 15.30           | 0.80     | 0.333    | 1.148                | 0.110           | 1.171     |
| 6           | 17.70        | 16.90           | 0.80     | 0.333    | 1.148                | 0.110           | 1.171     |
| 7           | 16.60        | 16.90           | -0.30    | 0.333    | -0.430               | 0.015           | -0.410    |
| 8           | 16.40        | 16.90           | -0.50    | 0.333    | -0.717               | 0.043           | -0.696    |
| 9           | 8.80         | 9.40            | -0.60    | 0.500    | -0.994               | 0.165           | -0.993    |
| 10          | 10.00        | 9.40            | 0.60     | 0.500    | 0.994                | 0.165           | 0.993     |
| 11          | 10.00        | 10.50           | -0.50    | 0.333    | -0.717               | 0.043           | -0.696    |
| 12          | 9.70         | 10.50           | -0.80    | 0.333    | -1.148               | 0.110           | -1.171    |
| 13          | 11.80        | 10.50           | 1.30     | 0.333    | 1.865                | 0.290           | 2.245     |
| 14          | 16.80        | 16.40           | 0.40     | 0.500    | 0.663                | 0.073           | 0.641     |
| 15          | 16.00        | 16.40           | -0.40    | 0.500    | -0.663               | 0.073           | -0.641    |

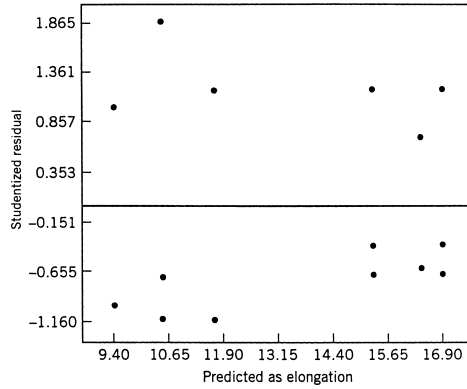
studentized residuals (in contrast to the ordinary least squares residuals) from mixture experiments, because the points in mixture designs can have substantial differences in their leverage values. Studentized residuals account for leverage through the factor  $1-h_{ii}$  that appears in the denominator (see Eq. 2.49 in Chapter 2).

Figure 12.8 presents a contour plot of elongation and a three-dimensional response surface plot. These plots are helpful in interpreting the results. From examining the figure, we note that if maximum elongation is desired, a blend of components 1 and 3 should be chosen consisting of about two-thirds component 3 and one-third component 1.

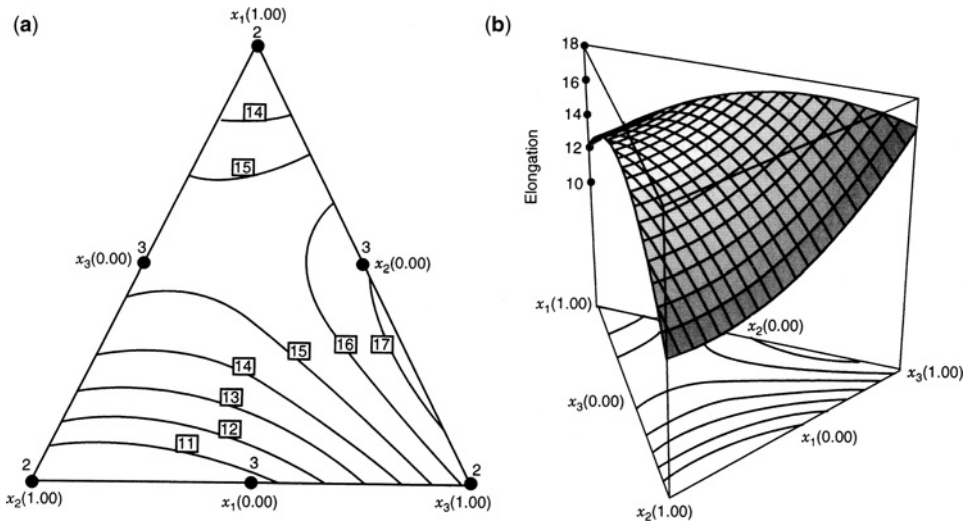




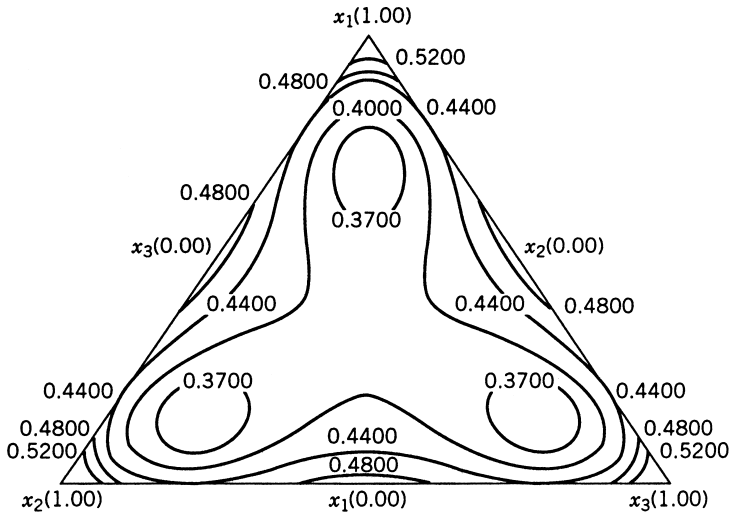
**Figure 12.6** Normal probability plot of studentized residuals for the quadratic model of yarn elongation.



**Figure 12.7** Plot of studentized residuals versus predicted yarn elongation.



**Figure 12.8** (a) Contour plot of predicted yarn elongation. (b) Three-dimensional surface plot. The solid dots on the contour plot are the design points.



**Figure 12.9** Contours of constant standard deviation of predicted response  $\sqrt{\text{Var}[\hat{y}(\mathbf{x})]}$  for the quadratic mixture model for the yarn elongation data.

Figure 12.9 presents the contours of constant standard deviation of predicted yarn elongation  $\left(\sqrt{\text{Var}[\hat{y}(\mathbf{x})]}\right)$  over the complete mixture space. This is related to the estimated prediction variance (EPV) defined in Chapter 8. Notice that the quality of prediction is better, in a variance sense, at points in the interior of the simplex than on the edges or at the vertices. However, one should be very careful about making predictions at points in the interior, because no experimental trials were performed there. The  $\{3, 2\}$  simplex lattice is a **boundary point design**; that is, all of the design points are either binary mixtures or pure mixtures. In this application the engineers were not interested in the ternary blends, so the lack of interior points is not a major concern. When one is interested in prediction in the interior, it is highly desirable to augment simplex-type designs with interior design points. We will return to this again in Section 12.2.3.

JMP also has excellent capability to design and analyze mixture experiments. Table 12.4 contains the JMP analysis for the yarn elongation experiment in Example 12.1 a plot of actual versus predicted values is at the top of the output display, followed by a set of summary statistics, and an analysis of variance for the second-order model. The JMP ANOVA output matches that from Design-Expert in Table 12.3. JMP does compute and display  $t$ -statistic for the individual linear blending terms while Design-Expert does not. At the bottom of the table is the JMP prediction profiler. We have set mixture component  $x_2 = 0$  in the profiler so that all values of the other two components are formulation of the binary mixture containing  $x_1$  and  $x_3$ . The formulation shown is  $x_1 = 0.315$  and  $x_3 = 0.685$ , yielding an elongation of 17.38.

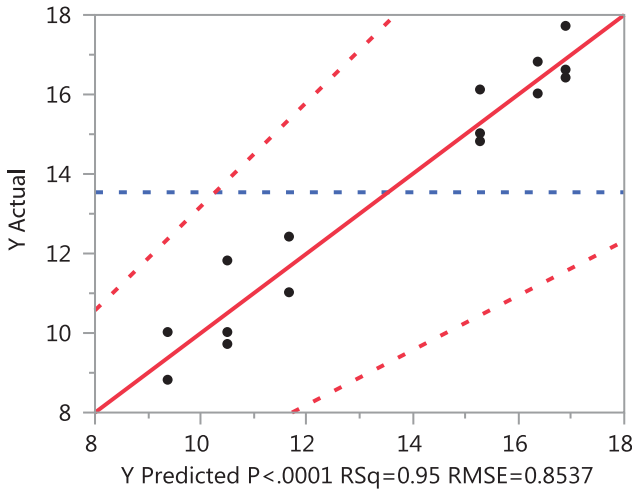
### 12.2.2 The Simplex-Centroid Design and Its Associated Polynomial

A  $q$ -component simplex-centroid design consists of  $2^q - 1$  distinct design points. These design points are the  $q$  permutations of  $(1, 0, 0, \dots, 0)$  or single-component blends, the  $\binom{q}{2}$  permutations of  $\left(\frac{1}{2}, \frac{1}{2}, 0, \dots, 0\right)$  or all binary mixtures, the  $\binom{q}{3}$  permutations of

**TABLE 12.4 JMP Output for the Yarn Elongation Experiment in Example 12.1**

**Response Y**

**Actual by Predicted Plot**



**Summary of Fit**

|                               |          |
|-------------------------------|----------|
| RSquare                       | 0.951356 |
| RSquare Adj                   | 0.924331 |
| Root Mean Square Error        | 0.85375  |
| Mean of Response              | 13.54    |
| Observations (or Sum Weights) | 15       |

**Analysis of Variance**

| Source   | DF | Sum of Squares | Mean Square | F-Ratio            |
|----------|----|----------------|-------------|--------------------|
| Model    | 5  | 128.29600      | 25.6592     | 35.2032            |
| Error    | 9  | 6.56000        | 0.7289      | <b>Prob &gt; F</b> |
| C. Total | 14 | 134.85600      |             | <.0001*            |

Tested against reduced model:  $Y = \text{mean}$

**Parameter Estimates**

| Term  | Estimate | Std Error | t-Ratio | Prob >  t |
|-------|----------|-----------|---------|-----------|
| X1    | 11.7     | 0.603692  | 19.38   | <.0001*   |
| X2    | 9.4      | 0.603692  | 15.57   | <.0001*   |
| X3    | 16.4     | 0.603692  | 27.17   | <.0001*   |
| X1*X2 | 19       | 2.608249  | 7.28    | <.0001*   |
| X1*X3 | 11.4     | 2.608249  | 4.37    | 0.0018*   |
| X2*X3 | -9.6     | 2.608249  | -3.68   | 0.0051*   |

(continued)

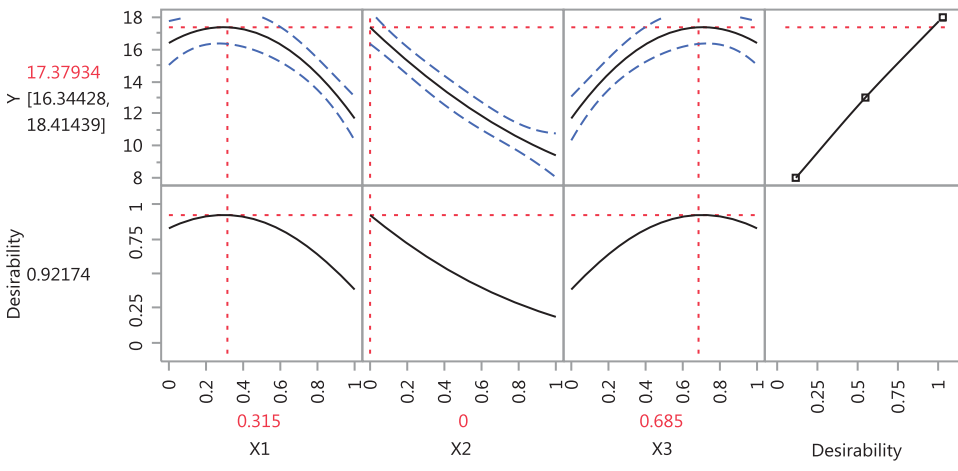
TABLE 12.4 (Continued)

| Effect Tests |        |    |                |          |          |
|--------------|--------|----|----------------|----------|----------|
| Source       | Nparam | DF | Sum of Squares | F-Ratio  | Prob > F |
| X1           | 1      | 1  | 273.78000      | 375.6128 | <.0001*  |
| X2           | 1      | 1  | 176.72000      | 242.4512 | <.0001*  |
| X3           | 1      | 1  | 537.92000      | 738.0000 | <.0001*  |
| X1*X2        | 1      | 1  | 38.67857       | 53.0651  | <.0001*  |
| X1*X3        | 1      | 1  | 13.92429       | 19.1034  | 0.0018*  |
| X2*X3        | 1      | 1  | 9.87429        | 13.5470  | 0.0051*  |

Sorted Parameter Estimates

| Term  | Estimate | Std Error | t-Ratio | t-Ratio | Prob >  t |
|-------|----------|-----------|---------|---------|-----------|
| X3    | 16.4     | 0.603692  | 27.17   |         | <.0001*   |
| X1    | 11.7     | 0.603692  | 19.38   |         | <.0001*   |
| X2    | 9.4      | 0.603692  | 15.57   |         | <.0001*   |
| X1*X2 | 19       | 2.608249  | 7.28    |         | <.0001*   |
| X1*X3 | 11.4     | 2.608249  | 4.37    |         | 0.0018*   |
| X2*X3 | -9.6     | 2.608249  | -3.68   |         | 0.0051*   |

Prediction Profiler



$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, \dots, 0)$ , and so forth, and the overall centroid  $(1/q, 1/q, \dots, 1/q)$ . Figure 12.10 shows the simplex-centroid designs for  $q = 3$  and  $q = 4$  components. Note that the design points are located at the centroid of the  $(q - 1)$ -dimensional simplex and at the centroids of all the lower-dimensional simplices contained within the  $(q - 1)$ -dimensional simplex.

The design points in the simplex-centroid design support the polynomial

$$\begin{aligned}
 E(y) = & \sum_{i=1}^q \beta_i x_i + \sum_{i < j=2}^q \sum \beta_{ij} x_i x_j + \sum_{i < j < k=3}^q \sum \sum \beta_{ijk} x_i x_j x_k \\
 & + \dots + \beta_{12\dots q} x_1 x_2 \dots x_q
 \end{aligned}
 \tag{12.11}$$

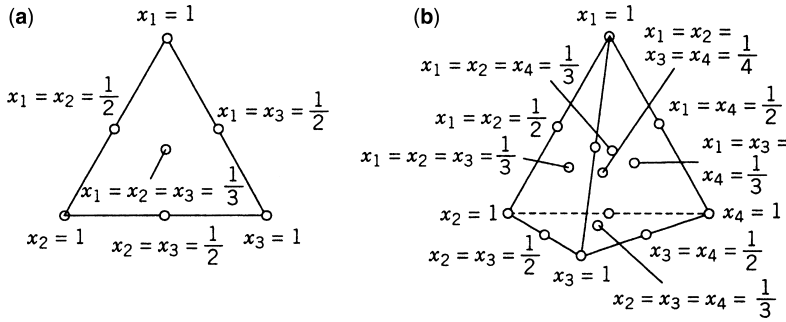


Figure 12.10 Simplex-centroid designs with three and four components. (a)  $q = 3$ . (b)  $q = 4$ .

which is a  $q$ th-order mixture polynomial. For  $q = 3$  components, this model is

$$E(y) = \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{12}x_1x_2 + \beta_{13}x_1x_3 + \beta_{23}x_2x_3 + \beta_{123}x_1x_2x_3$$

which is the special cubic polynomial from Equation 12.10. For  $q = 4$  components, the model is

$$E(y) = \sum_{i=1}^4 \beta_i x_i + \sum_{i < j=2}^4 \sum \beta_{ij} x_i x_j + \sum_{i < j < k=3}^4 \sum \sum \beta_{ijk} x_i x_j x_k + \beta_{1234} x_1 x_2 x_3 x_4$$

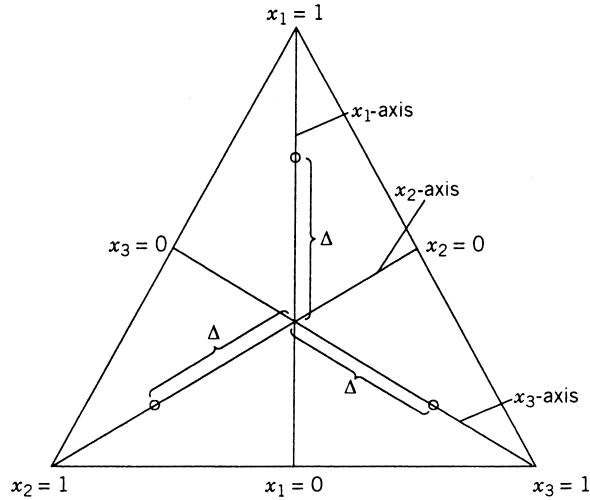
or the special cubic model with an additional quartic term. Because they are relatively efficient designs for fitting the special cubic model, simplex-centroid designs are often used when the experimenter thinks that some cubic terms may be necessary in the final model.

### 12.2.3 Augmentation of Simplex Designs with Axial Runs

The standard simplex-lattice and simplex-centroid designs are **boundary point designs**; that is, with the exception of the overall centroid, all the design points are on the boundaries of the simplex. For example, consider the  $\{3, 3\}$  lattice design shown in Fig. 12.3b. This design has ten points: the three vertices, the six points that are at the thirds of the edges, and the overall centroid. Thus there are three points that tell us something about the pure blends, six points that provide information about binary blends, and only one point that tells us anything about complete mixtures. We could say that the distribution of information about these different types of mixtures is 3:6:1.

If one is interested in making predictions about the properties of complete mixtures, it would be highly desirable to have more runs in the interior of the simplex. We recommend augmenting the usual simplex designs with **axial runs** and the overall centroid (if the centroid is not already a design point).

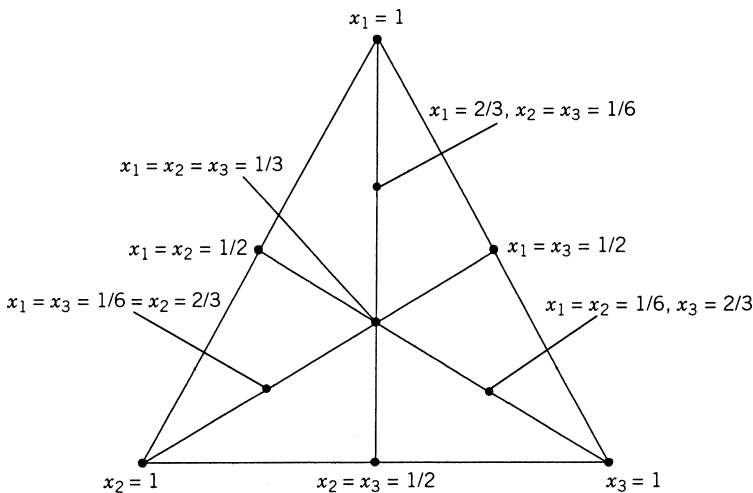
We define the **axis of component  $i$**  as the line or ray extending from the base point  $x_i = 0, x_j = 1/(q - 1)$  for all  $j \neq i$  to the opposite vertex where  $x_i = 1, x_j = 0$  for all  $j \neq i$ . Figure 12.11 shows the axes for components 1, 2, and 3 in a three-component system. Notice that the base point will always lie at the centroid of the  $(q - 2)$ -dimensional boundary of



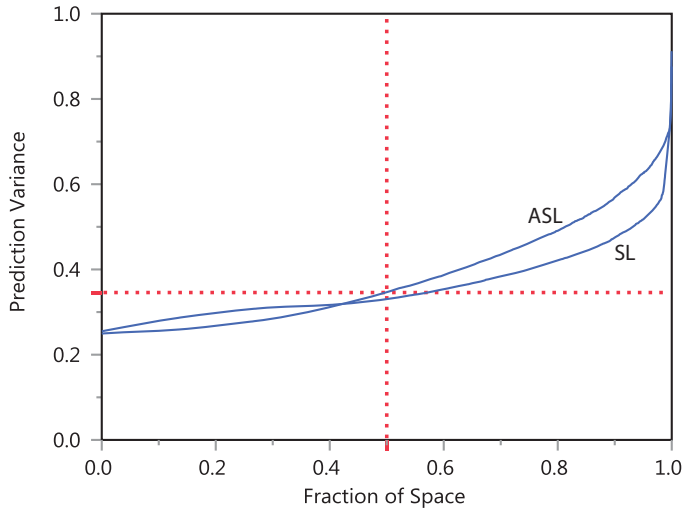
**Figure 12.11** The axes for components  $x_1$ ,  $x_2$ , and  $x_3$ .

the simplex that is opposite the vertex  $x_i = 1, x_j = 0$  for all  $j \neq i$ . [The boundary is sometimes called a  $(q - 2)$ -flat.] The length of the component axis is one unit. **Axial points** are positioned along the component axes a distance  $\Delta$  from the centroid. The maximum value for  $\Delta$  is  $(q - 1)/q$ . We recommend that axial runs be placed midway between the centroid of the simplex and each vertex, so that  $\Delta = (q - 1)/2q$ . Sometimes these points are called **axial check blends**, because a fairly common practice is to exclude them when fitting the preliminary mixture model, and then use the responses at these axial points to check the adequacy of the fit of the preliminary model.

Figure 12.12 shows the  $\{3, 2\}$  simplex-lattice design augmented with the axial points. This design has 10 points, with four of these points in the interior of the simplex. Thus the



**Figure 12.12** An augmented simplex-lattice design.



**Figure 12.13** Fraction of design space plot for the augmented (3, 2) simplex lattice (ASL) and the (3, 3) simplex lattice (SL) design for fitting a quadratic mixture model.

distribution of information in this augmented simplex lattice is 3:3:4. Contrast this to the 10-point  $\{3, 3\}$  simplex lattice, for which the distribution of information is 3:6:1. Figure 12.13 is the fraction of design space plot from JMP comparing these two designs. The prediction variance for fitting a quadratic mixture model for both designs is very similar. However, the augmented (3, 2) simplex lattice design has slightly smaller prediction variance than the (3, 3) simplex lattice design over approximately 45% of the design space. Cornell (1986) presents an extensive comparison of these two designs. He notes that the  $\{3, 3\}$  simplex lattice will support fitting the full cubic model, while the augmented simplex lattice will not; however, the augmented simplex lattice will allow the experimenter to fit the special cubic model or to add special quadratic terms such as  $\beta_{1233}x_1x_2x_3^2$  to the quadratic model. The augmented simplex lattice is superior for studying the response of complete mixtures in the sense that it can detect and model curvature in the interior of the triangle that cannot be accounted for by the terms in the full cubic model. The augmented simplex lattice has more power for detecting lack of fit than does the  $\{3, 3\}$  lattice. This is particularly useful when the experimenter is unsure about the proper model to use and also plans to **sequentially** build a model by starting with a simple polynomial (perhaps first order), test the model for lack of fit, then augment the model with higher-order terms, test the new model for lack of fit, and so forth.

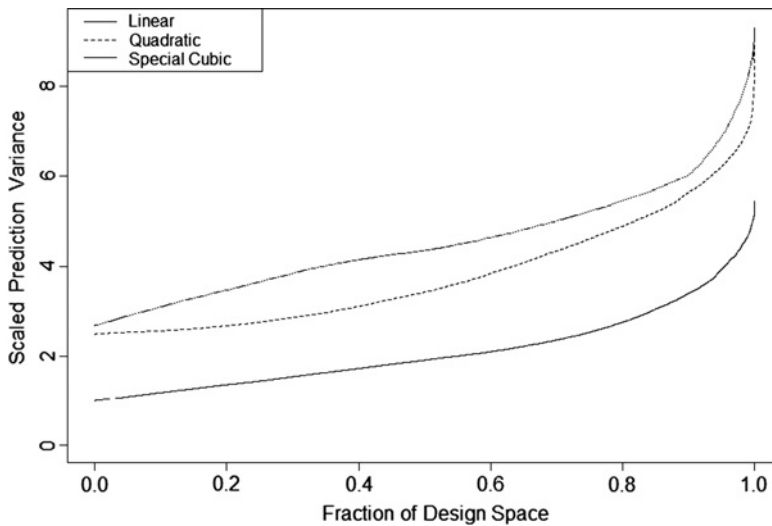
**Example 12.2 Etching Wafers** Wet chemical etching is often performed on the backs of silicon wafers prior to metallization in the semiconductor industry. The etching solution is a mixture of three different acids: A, B, and C. The experimenters wish to study how the composition of this mixture affects the etch rate. The experimenters think that a quadratic mixture model might be appropriate for modeling the relationship between etch rate and the proportions of different acids. However, there is some uncertainty about the nature of this relationship, so linear and special cubic models will also be considered. The experimenters selected the augmented simplex-lattice design shown in Fig. 12.12 for the use in this study, since it is a good choice for the quadratic model, but also performs well for the linear and

**TABLE 12.5** Augmented Simplex-Lattice Design for the Etch Rate Experiment

| Design Point | $x_1$ (Acid A) | $x_2$ (Acid B) | $x_3$ (Acid C) | Etch Rate<br>$y$ ( $\text{\AA}/\text{min}$ ) |
|--------------|----------------|----------------|----------------|--|
| 1            | 1              | 0              | 0              | 540, 560                                     |
| 2            | 0              | 1              | 0              | 330, 350                                     |
| 3            | 0              | 0              | 1              | 295, 260                                     |
| 4            | $\frac{1}{2}$  | $\frac{1}{2}$  | 0              | 610  |
| 5            | 0              | $\frac{1}{2}$  | $\frac{1}{2}$  | 330  |
| 6            | $\frac{1}{2}$  | 0              | $\frac{1}{2}$  | 425  |
| 7            | $\frac{2}{3}$  | $\frac{1}{6}$  | $\frac{1}{6}$  | 710  |
| 8            | $\frac{1}{6}$  | $\frac{2}{3}$  | $\frac{1}{6}$  | 640  |
| 9            | $\frac{1}{6}$  | $\frac{1}{6}$  | $\frac{2}{3}$  | 460  |
| 10           | $\frac{1}{3}$  | $\frac{1}{3}$  | $\frac{1}{3}$  | 800, 850                                     |

special cubic cases if either of these turn out to be appropriate. They decided to replicate the vertices and the overall centroid so that an internal estimate of error could be obtained. Thus the complete experimental design, shown in Table 12.5 along with the response, consists of 14 runs at 10 distinct locations.

Figure 12.14 shows the fraction of design space (FDS) plot for the scaled prediction variance (SPV) for the three possible models. While previously we used the FDS plots for comparing different designs with the same assumed model, here we compare what quality of prediction we can expect depending on which final model is selected for a single design.



**Figure 12.14** Fraction of design space plot for the partially replicated augmented simplex-lattice design for the wafer etching example.



**TABLE 12.6 Sequential Model Fitting for the Etch Rate Data**

| <i>Sequential Model Sum of Squares</i> |                |                    |             |         |                 |
|--|----------------|--------------------|-------------|---------|-----------------|
| Source                                 | Sum of Squares | Degrees of Freedom | Mean Square | F-Value | Probability > F |
| Mean                                   | 3,661,828.6    | 1                  | 3,661,828.6 |         |                 |
| Linear                                 | 133,755.0      | 2                  | 66,877.5    | 2.13    | 0.165           |
| Quadratic                              | 229,364.9      | 3                  | 76,455.0    | 5.29    | 0.027           |
| Special cubic                          | 107,877.8      | 1                  | 107,877.8   | 96.52   | 0.001           |
| Full cubic                             | 4,240.9        | 2                  | 2,120.4     | 2.96    | 0.142           |
| Residual                               | 3,582.9        | 5                  | 716.5       |         |                 |
| Total                                  | 4,127,850.0    | 14                 |             |         |                 |

| <i>Lack-of-Fit Tests</i> |                |                    |             |         |                 |
|--------------------------|----------------|--------------------|-------------|---------|-----------------|
| Model                    | Sum of Squares | Degrees of Freedom | Mean Square | F-Value | Probability > F |
| Linear                   | 342,803.9      | 7                  | 48,972.0    | 86.58   | 0.001           |
| Quadratic                | 113,439.1      | 4                  | 28,359.8    | 50.14   | 0.001           |
| Special cubic            | 5,561.3        | 3                  | 1,853.8     | 3.28    | 0.141           |
| Full cubic               | 1,320.4        | 1                  | 1,320.4     | 2.33    | 0.201           |
| Pure Error               | 2,262.5        | 4                  | 565.6       |         |                 |

| <i>ANOVA Summary Statistics of Model Fit</i> |                 |                             |                         |           |                    |           |
|--|-----------------|-----------------------------|-------------------------|-----------|--------------------|-----------|
| Source                                       | Unaliated Terms | Residual Degrees of Freedom | Root Mean Squared Error | R-Squared | Adjusted R-Squared | PRESS     |
| Linear                                       | 3               | 11                          | 177.1                   | 0.2793    | 0.1483             | 516,539.9 |
| Quadratic                                    | 6               | 8                           | 120.3                   | 0.7584    | 0.6073             | 812,042.5 |
| Special cubic                                | 7               | 7                           | 33.4                    | 0.9837    | 0.9697             | 80,312.0  |
| Full cubic                                   | 9               | 5                           | 26.8                    | 0.9925    | 0.9805             | 186,374.0 |

See Ozol-Godfrey et al. (2005) for more details on assessing model robustness. As we add additional terms to the model, the SPV must necessarily increase for a given design, hence the linear model has the lowest SPV throughout the region. However, if the true underlying relationship between etch rate and mixture ingredients is quadratic or special cubic, then the linear model will have large bias and will miss important features in the relationship. Hence, we wish to select the simplest model which appropriately summarizes the true relationship.

Table 12.6 presents the results of fitting linear, quadratic, special cubic, and full cubic models to the data obtained in this experiment (the computations were performed using Design-Expert). The sequential *F*-tests in this table indicate that the contribution of the quadratic terms to the model (over the linear terms) is significant, as is the contribution of the special cubic term  $\beta_{123x_1x_2x_3}$  (over the linear and quadratic terms). Furthermore, the quadratic model displays significant lack of fit, whereas the special cubic model does not; and, in addition, the special cubic model has a smaller error mean square, a larger adjusted  $R^2$ , and a smaller value of the PRESS statistic than does the quadratic

model. Terms in the full cubic model are aliased and should be ignored. On the basis of this analysis, the experimenters selected the special cubic model. Table 12.7 presents the results of fitting the special cubic model. The upper portion of this table is the overall analysis of variance for this model. The  $F$ -test for the model is used to test the hypotheses

$$H_0: \beta_1 = \beta_2 = \beta_3 = \beta, \beta_{12} = \beta_{13} = \beta_{23} = \beta_{123} = 0$$

$$H_1: \text{At least one equality is false}$$

Because  $F_0 = 70.23$  ( $P < 0.001$ ), the null hypothesis is rejected. The  $t$ -statistics on the individual model terms indicate that two of the quadratic terms,  $x_1x_3$  and  $x_2x_3$ , could be deleted; however, the experimenters decided to retain them in order to keep the model hierarchical. The  $F$ -statistic for lack of fit is testing for contributions of terms beyond the special cubic. This design could support three special quartic terms. The lack-of-fit test indicates that these terms are unnecessary. The diagnostic information at the bottom of Table 12.7 is generally satisfactory. The center-of-edge points in this design have relatively large leverage values (0.912), and this is responsible in part for the three large values of Cook's distance (1.453, 1.204, and 4.862) reported in the table.

The fitted special cubic model is

$$\begin{aligned} \hat{y} = & 550.2x_1 + 344.7x_2 + 268.3x_3 + 689.5x_1x_2 - 9.0x_1x_3 \\ & + 58.1x_2x_3 + 9243.3x_1x_2x_3 \end{aligned}$$

Figures 12.15 and 12.16 present the normal probability plot of the studentized residuals and a plot of the studentized residuals versus the predicted values, respectively, for this model. The plots are satisfactory, so we conclude that the special cubic model is adequate to describe the etch rate response surface.

Figure 12.17 presents plots of the etch rate response surface. In this process it is important to obtain an etch rate that is at least  $750 \text{ \AA}/\text{min}$ . The contour plot in Fig. 12.17a indicates that there are several formulations that will meet that requirement.

We would like to emphasize an important point from Example 12.2. Notice that the experimenters originally intended to fit a quadratic mixture model; however, after the experiment had been performed, it was clear that a higher-order model (the special cubic) would be required to adequately model the response. This happens frequently in mixture experiments: The order of the required model is higher than the one initially planned for by the experimenters. If the experimenter designs the experiment without some additional runs to measure lack of fit (and fit higher-order terms, if necessary) and the proposed model is inadequate, then he or she is in trouble. For that reason, we recommend that the experimenter should, whenever possible, make 8–10 additional runs beyond the minimum required to fit the model, with about half of these runs chosen to be replicates of some points in the design and the others chosen as new distinct points so that the lack of fit of the model can be investigated. In this example, the tentative model was the quadratic, which in three components is a six-parameter model. The use of the augmented simplex lattice automatically provided four additional distinct runs. Then the experimenters decided to replicate four runs to give an internal estimate of error. Whenever it is possible to do so, we strongly recommend this design strategy.

**TABLE 12.7 Special Cubic Model for the Etch Rate Experiment**

| Source                   | Sum of Squares       | Degrees of Freedom | Mean Square            | F-Value                                 | Probability > F          |      |
|--------------------------|----------------------|--------------------|------------------------|---|--------------------------|------|
| Model                    | 470997.6             | 6                  | 78499.6                | 70.23                                   | <0.001                   |      |
| Residual                 | 7823.8               | 7                  | 1117.7                 |   |                          |      |
| Lack of fit              | 5561.3               | 3                  | 1853.8                 | 3.28                                    | 0.141                    |      |
| Pure error               | 2262.5               | 4                  | 565.6                  |   |                          |      |
| Corrected total          | 478821.4             | 13                 |                        |   |                          |      |
| Root mean squared error  | 33.4                 |                    | <i>R</i> -Squared      | 0.9837                                  | PRESS = 80,312.0         |      |
| Dependent Mean           | 511.4                |                    | Adj <i>R</i> -Squared  | 0.9697                                  |                          |      |
| Coefficient of variation | 6.54%                |                    | Pred <i>R</i> -Squared | 0.8323                                  |                          |      |
| Component                | Coefficient Estimate | Degrees of Freedom | Standard Error         | <i>t</i> for $H_0$ :<br>Coefficient = 0 | Probability >   <i>t</i> | VIF  |
| $x_1$                    | 550.2                | 1                  | 23.2                   |   | Not applicable           | 1.55 |
| $x_2$                    | 344.7                | 1                  | 23.2                   |   | Not applicable           | 1.55 |
| $x_3$                    | 268.3                | 1                  | 23.2                   |   | Not applicable           | 1.55 |
| $x_1x_2$                 | 689.5                | 1                  | 146.5                  | 4.71                                    | 0.002                    | 2.16 |
| $x_1x_3$                 | -9.0                 | 1                  | 146.5                  | -0.06                                   | 0.953                    | 2.16 |
| $x_2x_3$                 | 58.1                 | 1                  | 146.5                  | 0.40                                    | 0.703                    | 2.16 |
| $x_1x_2x_3$              | 9243.3               | 1                  | 940.9                  | 9.82                                    | <0.001                   | 2.99 |

(continued)

TABLE 12.7 (Continued)

| Observation | Actual Value | Predicted Value | Residual | $h_{ii}$ | Student Residual | Cook's Distance | R-Student |
|-------------|--------------|-----------------|----------|----------|------------------|-----------------|-----------|
| 1           | 540.0        | 550.2           | -10.2    | 0.483    | -0.42            | 0.024           | -0.40     |
| 2           | 560.0        | 550.2           | 9.8      | 0.483    | 0.41             | 0.022           | 0.38      |
| 3           | 330.0        | 344.7           | -14.7    | 0.483    | -0.61            | 0.050           | -0.58     |
| 4           | 350.0        | 344.7           | 5.3      | 0.483    | 0.22             | 0.006           | 0.20      |
| 5           | 295.0        | 268.3           | 26.7     | 0.483    | 1.11             | 0.164           | 1.13      |
| 6           | 260.0        | 268.3           | -8.3     | 0.483    | -0.34            | 0.016           | -0.32     |
| 7           | 610.0        | 619.8           | -9.8     | 0.912    | -0.99            | 1.453           | -0.99     |
| 8           | 425.0        | 407.0           | 18.0     | 0.912    | 1.81             | 4.862           | 2.31      |
| 9           | 330.0        | 321.0           | 9.0      | 0.912    | 0.90             | 1.204           | 0.89      |
| 10          | 800.0        | 812.2           | -12.2    | 0.375    | -0.46            | 0.018           | -0.43     |
| 11          | 850.0        | 812.2           | 37.8     | 0.375    | 1.43             | 0.176           | 1.58      |
| 12          | 710.0        | 717.4           | -7.4     | 0.206    | -0.25            | 0.002           | -0.23     |
| 13          | 640.0        | 620.2           | 19.8     | 0.206    | 0.66             | 0.016           | 0.64      |
| 14          | 460.0        | 523.8           | -63.8    | 0.206    | -2.14            | 0.170           | -3.38     |

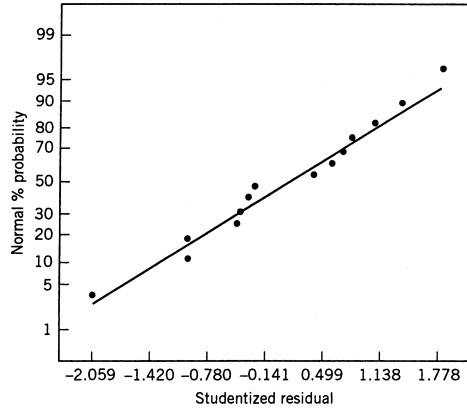


Figure 12.15 Normal probability plot of the studentized residuals.

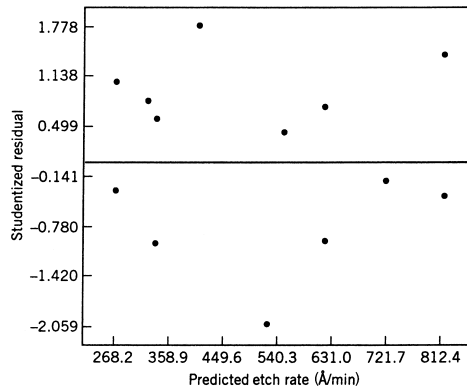


Figure 12.16 Plot of studentized residuals versus predicted etch rate.

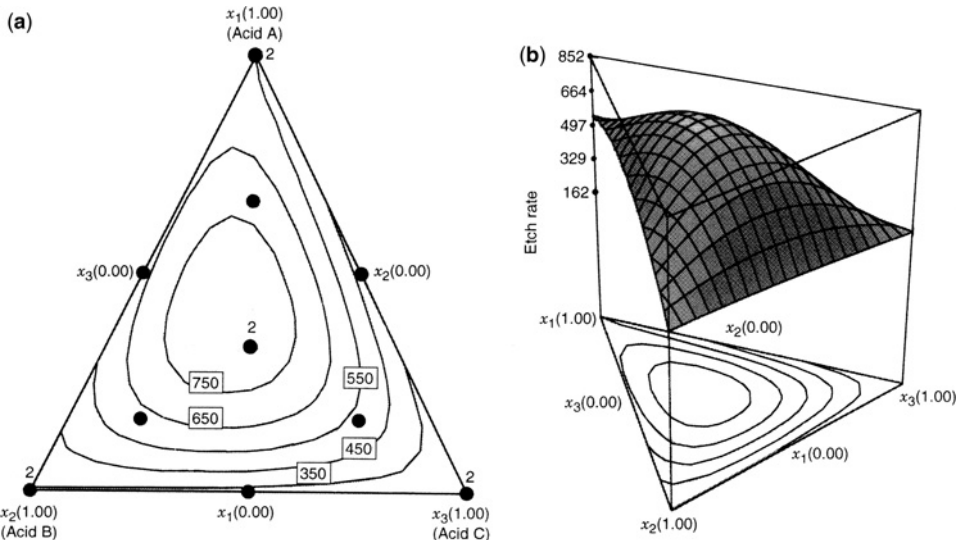


Figure 12.17 Etch rate response surface. (a) Contour plot. (b) Three-dimensional surface plot. The solid dots in the contour plot are the design points.

### 12.3 RESPONSE TRACE PLOTS

The **response trace** is a plot of the estimated response values as we move away from the centroid of the simplex and along the component axes. More generally, we may think of the centroid as a **reference mixture** with component proportions given by  $r_i = 1/q_i$ ,  $i = 1, 2, \dots, q$  (in Chapter 12 we will discuss other choices for the proportions in the reference blend). Consider the  $i$ th component, and suppose we move away from the centroid by changing the proportion of this component by an amount  $\Delta_i$  (note that we could make  $\Delta_i$  either positive or negative). Along the  $i$ th axis, as the value of  $x_i$  either increases or decreases, the values of the other component proportions  $x_j, j \neq i$ , either decrease or increase, but the relative proportions for these other components remain the same.

To illustrate, consider a  $q = 3$  component mixture, so that  $r_i = \frac{1}{3}$ ,  $i = 1, 2, 3$ . Suppose that component 1 is increased by  $\Delta_1 = \frac{1}{6}$  so that now  $x_1 = r_1 + \Delta_1 = \frac{1}{3} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$ . The new proportions for components 2 and 3 are  $x_2 = x_3 = \frac{1}{4}$ . Thus the ratio of  $x_2$  to  $x_3$  is identical at the centroid ( $x_2/x_3 = \frac{1/3}{1/3} = 1$ ) and at the new point along the  $x_1$ -axis, because at this point  $x_2/x_3 = \frac{1/4}{1/4} = 1$ .

We may give a general equation for these results. If  $x_i$  is changed by an amount  $\Delta_i$  from the reference mixture, then the new proportions are

$$\begin{aligned} x_i &= r_i + \Delta_i \\ x_j &= r_j - \frac{\Delta_i r_j}{1 - r_i}, \quad j \neq i \end{aligned} \quad (12.12)$$

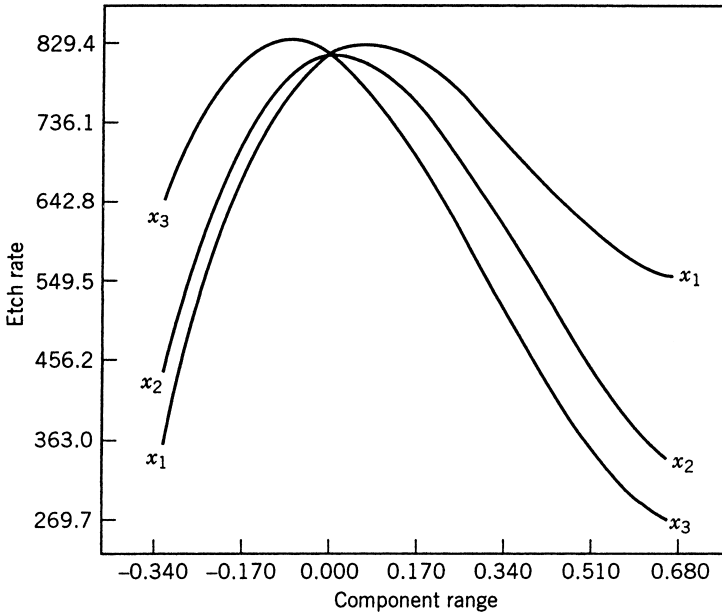
To construct a response surface trace plot, we choose some number of blends along each component axis, calculate the coordinates of each mixture using Equation (12.12), and then substitute these coordinates into the fitted mixture model to obtain the predicted values of the response. Then these predicted responses are plotted against the changes made in the  $x_i$ ,  $i = 1, 2, \dots, q$ . There will be  $q$  of these curves, and it is customary to plot them on the same graph.

To illustrate, Fig. 12.18 is the response trace plot for the etch rate experiment described in Example 12.2. The vertical axis is the predicted etch rate, and the horizontal axis is the incremental change  $\Delta_i$  made in each component. The reference mixture, which is the overall centroid, is shown as the point 0.000 on the horizontal axis. By definition, all trace plots must have the same etch rate at 0, since this is the reference mixture. Notice the strong nonlinear effect of all three acids on the etch rate. Also, the etch rate increases slightly and then decreases as we move along the  $x_1$ -axis from the centroid toward the  $x_1$ -vertex, and it increases as we move away from the centroid along the  $x_3$ -axis toward the base point  $x_3 = 0, x_1 = x_2 = \frac{1}{2}$ .

In this example, the etch rate is very sensitive to changes in all three component proportions. If one or more of the response traces is a horizontal line, this indicates that these components have little effect on the response; that is, we have discovered ingredients that are **inactive**. This can potentially have important economic consequences.

### 12.4 REPARAMETERIZING CANONICAL MIXTURE MODELS TO CONTAIN A CONSTANT TERM ( $\beta_0$ )

It is occasionally convenient to reparameterize the Scheffé mixture polynomial so that it contains a constant term, or intercept,  $\beta_0$ . This can be easily done by simply deleting one



**Figure 12.18** Response trace plot of predicted etch rate along the component axes  $x_1$ ,  $x_2$ , and  $x_3$ .

of the  $\beta_i x_i$  terms, say  $\beta_q x_q$ . Thus, instead of the linear mixture model

$$E(y) = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_q x_q \tag{12.13}$$

we would write

$$E(y) = \beta_0 + \beta_1^* x_1 + \beta_2^* x_2 + \dots + \beta_{q-1}^* x_{q-1} \tag{12.14}$$

and instead of the quadratic model

$$E(y) = \sum_{i=1}^q \beta_i x_i + \sum_{i < j=2}^q \beta_{ij} x_i x_j \tag{12.15}$$

we would write

$$E(y) = \beta_0 + \sum_{i=1}^{q-1} \beta_i^* x_i + \sum_{i < j=2}^q \beta_{ij} x_i x_j \tag{12.16}$$

Note that only the parameters in the linear blending portion of the canonical model are affected. Specifically, the intercept  $\beta_0$  takes on the role of the parameter  $\beta_q$  corresponding to the deleted component  $x_q$ , and the remaining parameters  $\beta_i^*$  represent the differences  $\beta_i^* = \beta_i - \beta_q$ ,  $i = 1, 2, \dots, q - 1$ .

We use the quadratic model for yarn elongation in Example 12.1 to illustrate. The fitted canonical model was

$$\hat{y} = 11.7x_1 + 9.4x_2 + 16.4x_3 + 19.0x_1x_2 + 11.4x_1x_3 - 9.6x_2x_3$$

Now fit the model in Equation 12.14 with  $q = 3$  by the method of least squares. This results in

$$\hat{y} = 16.4 - 4.7x_1 - 7.0x_2 + 19.0x_1x_2 + 11.4x_1x_3 - 9.6x_2x_3$$

That is,

$$b_0 = 16.4, \quad b_1^* = -4.7, \quad b_2^* = -7.0$$

Notice that

$$\begin{aligned} b_0 &= b_3 = 16.4 \\ b_1^* &= b_1 - b_3 = 11.7 - 16.4 = -4.7 \end{aligned}$$

and

$$b_2^* = b_2 - b_3 = 9.4 - 16.4 = -7.0$$

and that the estimates of the quadratic terms are the same in both models.

This reparameterization of the Scheffé polynomial can be useful in analyzing mixture data with standard multiple regression computer software. Most regression software packages will fit a model with the intercept term suppressed. This option will produce the correct estimates of the  $\beta$ 's in a canonical mixture model, but it will result in an incorrect value for the model or regression sum of squares, because the no-intercept option does not correct this sum of squares for the overall mean. As a result, the  $F$ -test,  $R^2$ , and the adjusted  $R^2$  will be incorrect. Reparameterizing the model as we have discussed will result in correct output for the regression or model sum of squares,  $R^2$ , and the adjusted  $R^2$ . Then the correct values of the linear blending coefficients can be calculated from

$$\begin{aligned} b_i &= b_i^* + b_0, \quad i = 1, 2, \dots, q-1 \\ b_q &= b_0 \end{aligned} \tag{12.17}$$

Another technique occasionally used in the analysis of mixture data is the **slack variable approach**. The approach differs from the previous one in that one of the mixture components is *completely eliminated* from the model. That is, if the component  $x_q$  is eliminated, then the first-order slack variable model is

$$\begin{aligned} E(y) &= \beta_0 + \beta_1^*x_1 + \beta_2^*x_2 + \dots + \beta_{q-1}^*x_{q-1} \\ &= \beta_0 + \sum_{i=1}^{q-1} \beta_i^*x_i \end{aligned} \tag{12.18}$$



**TABLE 12.8 Design-Expect Output for the Slack Variable Model of the Yarn Elongation Data**

| Response  | Elongation     |                |                |            |             |       |
|---|----------------|----------------|----------------|------------|-------------|-------|
| ***Component C is the Slack Variable.***            |                |                |                |            |             |       |
| ANOVA for Slack Mixture Quadratic Model             |                |                |                |            |             |       |
| Analysis of variance table [Partial sum of squares] |                |                |                |            |             |       |
| Source  | Sum of Squares | DF             | Mean Square    | F-Value    | Prob > F    |       |
| Model   | 128.30         | 5              | 25.66          | 32.20      | < 0.0001    |       |
| A   | 4.34           | 1              | 4.34           | 5.96       | 0.0373      |       |
| B   | 26.67          | 1              | 26.67          | 36.59      | 0.0002      |       |
| A <sup>2</sup>                                      | 13.92          | 1              | 13.92          | 19.10      | 0.0018      |       |
| B <sup>2</sup>                                      | 9.87           | 1              | 9.87           | 13.55      | 0.0051      |       |
| AB  | 12.33          | 1              | 12.33          | 16.91      | 0.0026      |       |
| Pure Error  | 6.56           | 9              | 0.73           |            |             |       |
| Cor Total   | 134.86         | 14             |                |            |             |       |
| Std. Dev.   | 0.85           | R-Squared      | 0.9514         |            |             |       |
| Mean  | 13.54          | Adj R-Squared  | 0.9243         |            |             |       |
| C.V.  | 6.31           | Pred R-Squared | 0.8643         |            |             |       |
| PRESS   | 18.30          | Adeq Precision | 13.890         |            |             |       |
| Component   | Estimate       | Coefficient DF | Standard Error | 95% CI Low | 95% CI High | VIF   |
| Intercept   | 16.40          | 1              | 0.60           | 15.03      | 17.77       |       |
| A-A   | -6.70          | 1              | 2.74           | 0.49       | 12.91       | 18.94 |
| B-B   | -16.60         | 1              | 2.74           | -22.81     | -10.39      | 18.94 |
| A <sup>2</sup>                                      | -11.40         | 1              | 2.61           | -17.30     | -5.50       | 14.54 |
| B <sup>2</sup>                                      | 9.60           | 1              | 2.61           | 3.70       | 15.50       | 14.54 |
| AB  | 17.20          | 1              | 4.18           | 7.74       | 26.66       | 3.60  |
| Final Equation in Terms of Pseudo Components:       |                |                |                |            |             |       |
| Elongation =  |                |                |                |            |             |       |
| +16.40  |                |                |                |            |             |       |
| +6.70*A   |                |                |                |            |             |       |
| -16.60*B  |                |                |                |            |             |       |
| -11.40*A <sup>2</sup>                               |                |                |                |            |             |       |
| +9.50*B <sup>2</sup>                                |                |                |                |            |             |       |
| +17.20*A*B  |                |                |                |            |             |       |

which is identical to Equation 12.13. However, the second-order slack variable model is

$$E(y) = \beta_0 + \sum_{i=1}^{q-1} \beta_i^* x_i + \sum_{i<j=2}^{q-1} \beta_{ij}^* x_i x_j + \sum_{i=1}^{q-1} \beta_{ii}^* x_i^2 \quad (12.19)$$

Note that this is a standard second-order response model in the  $q - 1$  mixture components. The slack variable version of the special cubic model is

$$\begin{aligned} E(y) = & \beta_0 + \sum_{i=1}^{q-1} \beta_i^* x_i + \sum_{i<j=2}^{q-1} \beta_{ij}^* x_i x_j + \sum_{i=1}^{q-1} \beta_{ii}^* x_i^2 \\ & + \sum_{i<j<k=2}^{q-1} \beta_{ijk}^* x_i x_j x_k + \sum_{i<j=2}^{q-1} \beta_{ijj}^* x_i x_j (x_i + x_j) \end{aligned} \quad (12.20)$$

It turns out that the coefficients in the slack variable model are just simple functions of the coefficients in the original Scheffé polynomials.

Some practitioners report using the slack variable model with some success particularly if one component does not impact the response, such as a filler ingredient like water in a shampoo. So some software packages support it. Design-Expert is one such package. Table 12.8 is the output from Design-Expert for the yarn elongation data from Example 12.1, using  $x_3$  (component C) as the slack variable. Notice that all the terms are significant in the model. Therefore, the slack variable model is completely equivalent to the original Scheffé quadratic that we fitted to these data in Example 12.1.

Cornell (2000) points out that if we find that some terms in the slack variable model are not significant and reduce the model (say by using stepwise regression techniques such as we discuss in Appendix A), then the reduced slack variable model will not necessarily be equivalent to the Scheffé model that was also reduced by eliminating nonsignificant terms. The contour plots and hence the final conclusions from the two model can be different. Furthermore, Cornell observes that the results one obtains can depend greatly on which variable is chosen as the slack variable. Fortunately, Design-Expert allows the user to specify the slack variable, so an experienced analyst can try different models and decide if the slack variable approach is really appropriate for his or her situation. Our own recommendation is to avoid indiscriminate use of the slack variable model in most problems and use the standard Scheffé polynomials unless there is some overriding reason to not do so. For an excellent discussion of the slack variable model and many more interesting details, refer to Cornell (2000).

## EXERCISES

- 12.1** Three different motor fuels can be blended to form a gasoline. The miles per gallon (MPG) performance of this gasoline is the response variable. An augmented simplex-lattice design was used to study the blending properties of these three fuels, and the results are shown in Table E12.1:
- Fit a linear mixture model to these data. Is the model adequate?
  - Fit a quadratic mixture model to these data. Is this model an improvement over the linear model?

**TABLE E12.1** Data for Exercise 12.1

| Design Point | Component Proportion |               |               | MPG $y$    |
|--------------|----------------------|---------------|---------------|------------|
|              | $x_1$                | $x_2$         | $x_3$         |            |
| 1            | 1                    | 0             | 0             | 24.5, 25.1 |
| 2            | 0                    | 1             | 0             | 24.8, 23.9 |
| 3            | 0                    | 0             | 1             | 22.7, 23.6 |
| 4            | $\frac{1}{2}$        | $\frac{1}{2}$ | 0             | 25.1       |
| 5            | $\frac{1}{2}$        | 0             | $\frac{1}{2}$ | 24.3       |
| 6            | 0                    | $\frac{1}{2}$ | $\frac{1}{2}$ | 23.5       |
| 7            | $\frac{1}{3}$        | $\frac{1}{3}$ | $\frac{1}{3}$ | 24.8, 24.1 |
| 8            | $\frac{2}{3}$        | $\frac{1}{6}$ | $\frac{1}{6}$ | 24.2       |
| 9            | $\frac{1}{6}$        | $\frac{2}{3}$ | $\frac{1}{6}$ | 23.9       |
| 10           | $\frac{1}{6}$        | $\frac{1}{6}$ | $\frac{2}{3}$ | 23.7       |

- (c) What model would you use to predict the MPG performance of the gasoline obtained by blending these three fuels?
- (d) Plot the response surface contours for the model in part (c). What blend would you use to maximize mileage performance?

**12.2** A mixture experiment was performed to determine if an artificial sweetener could be developed for a soft drink that would minimize the aftertaste. Four different sweeteners were evaluated, and a simplex-centroid design with  $q = 4$  components was used to evaluate the different blends. The response variable is aftertaste, with low values being desirable. The data are shown in Table E12.2.

- (a) Fit an appropriate mixture model to these data. Test the model for adequacy.
- (b) Construct the response trace plot for this experiment. Provide a practical interpretation of this plot.
- (c) Construct contour plots of predicted aftertaste that would be useful to a decisionmaker in choosing a blend of these four sweeteners that minimizes the aftertaste. What blend would you recommend?

**12.3** Consider the linear mixture model from Exercise 12.1a. Reparameterize this model to the form

$$E(y) = \beta_0 + \beta_1^* x_1 + \beta_2^* x_2$$

and fit this model using least squares. Verify that this model can be converted to the linear mixture model you found in Exercise 12.1a using the method discussed in Section 12.4.

**12.4** A thickening agent is added to a liquid soap product to control the viscosity. This agent usually makes up 10% of the soap by weight. There are three different salt compounds that are blended to make up the thickener. The product formulators have decided to investigate the combination of these ingredients using a mixture

TABLE E12.2 Data for Exercise 12.2

| Design Point | Component Proportions |               |               |               | Aftertaste<br>$y$ |
|--------------|-----------------------|---------------|---------------|---------------|-------------------|
|              | $x_1$                 | $x_2$         | $x_3$         | $x_4$         |                   |
| 1            | 1                     | 0             | 0             | 0             | 19                |
| 2            | 0                     | 1             | 0             | 0             | 8                 |
| 3            | 0                     | 0             | 1             | 0             | 15                |
| 4            | 0                     | 0             | 0             | 1             | 10                |
| 5            | $\frac{1}{2}$         | $\frac{1}{2}$ | 0             | 0             | 13                |
| 6            | $\frac{1}{2}$         | 0             | $\frac{1}{2}$ | 0             | 16                |
| 7            | $\frac{1}{2}$         | 0             | 0             | $\frac{1}{2}$ | 12                |
| 8            | 0                     | $\frac{1}{2}$ | $\frac{1}{2}$ | 0             | 11                |
| 9            | 0                     | $\frac{1}{2}$ | 0             | $\frac{1}{2}$ | 5                 |
| 10           | 0                     | 0             | $\frac{1}{2}$ | $\frac{1}{2}$ | 10                |
| 11           | $\frac{1}{3}$         | $\frac{1}{3}$ | $\frac{1}{3}$ | 0             | 14                |
| 12           | $\frac{1}{3}$         | $\frac{1}{3}$ | 0             | $\frac{1}{3}$ | 10                |
| 13           | $\frac{1}{3}$         | 0             | $\frac{1}{3}$ | $\frac{1}{3}$ | 14                |
| 14           | 0                     | $\frac{1}{3}$ | $\frac{1}{3}$ | $\frac{1}{3}$ | 8                 |
| 15           | $\frac{1}{4}$         | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | 12                |

experiment. Let  $x_1^*$ ,  $x_2^*$ , and  $x_3^*$  represent the actual proportions of the three salt compounds, with  $x_1^* + x_2^* + x_3^* = 0.1$ . Notice that if we let  $x_i = x_i^*/0.1$ , then the design points can be expressed as proportions  $0 \leq x_i \leq 1$  with the standard mixture constraint  $x_1 + x_2 + x_3 = 1$ . The experiments decided to use an augmented simplex-lattice design to study the viscosity of the thickener. The design and the viscosity data are shown in Table E12.3. The centroid was replicated three times.

- Fit linear and quadratic mixture models to these data. Test both models for lack of fit. Does either model seem adequate, based on this test?
- Select an appropriate mixture model for the viscosity data. Construct appropriate residual plots to check for model adequacy.
- Construct and interpret a response trace plot for this experiment.
- Construct a response surface contour plot for viscosity.
- Suppose that a desirable value for viscosity of this thickening agent is 900 centipoise. Is there a formulation that would produce the desired viscosity?

**12.5** One of the authors (DCM) was a partner in a vineyard and winery in Newberg, Oregon. The partners that operate this business are developing a new red wine that is a blend of three different varieties of grapes: Pinot Noir–Pommard, Pinot Noir–Wadenswil, and Gamay Noir. A three-component mixture design was used to study the blends of these three ingredients. The design is shown in Table E12.4 where  $x_1$  = proportion of Pinot Noir–Pommard,  $x_2$  = proportion of Pinot Noir–Wadenswil, and  $x_3$  = proportion of Gamay Noir in the product. Each of the 10 blends was tasted

TABLE E12.3 Data for Exercise 12.4

| Design Point | Component Proportions |               |               | Viscosity (centipoise) |
|--------------|-----------------------|---------------|---------------|------------------------|
|              | $x_1$                 | $x_2$         | $x_3$         |                        |
| 1            | 1                     | 0             | 0             | 410                    |
| 2            | 0                     | 1             | 0             | 880                    |
| 3            | 0                     | 0             | 1             | 1445                   |
| 4            | $\frac{1}{2}$         | $\frac{1}{2}$ | 0             | 683                    |
| 5            | $\frac{1}{2}$         | 0             | $\frac{1}{2}$ | 465                    |
| 6            | 0                     | $\frac{1}{2}$ | $\frac{1}{2}$ | 456                    |
| 7            | $\frac{2}{3}$         | $\frac{1}{6}$ | $\frac{1}{6}$ | 354                    |
| 8            | $\frac{1}{6}$         | $\frac{2}{3}$ | $\frac{1}{6}$ | 521                    |
| 9            | $\frac{1}{6}$         | $\frac{1}{6}$ | $\frac{2}{3}$ | 700                    |
| 10           | $\frac{1}{3}$         | $\frac{1}{3}$ | $\frac{1}{3}$ | 465                    |
| 11           | $\frac{1}{3}$         | $\frac{1}{3}$ | $\frac{1}{3}$ | 392                    |
| 12           | $\frac{1}{3}$         | $\frac{1}{3}$ | $\frac{1}{3}$ | 428                    |

by a panel of experts and ranked on the basis of the taste-test results, with rank 1 being the best, rank 2 the next best, and so forth.

- (a) Fit linear and quadratic mixture models to the averages of the rank data. Test both of these models for lack of fit. Does either model seem adequate?
- (b) Select an appropriate mixture model for the average rank response. Construct appropriate residual plots to check for model adequacy.

TABLE E12.4 Data for Exercise 11.5

| Design Point | Component Proportions |               |               | Taste-Test Ranks |     |     |     |
|--------------|-----------------------|---------------|---------------|------------------|-----|-----|-----|
|              | $x_1$                 | $x_2$         | $x_3$         | DCM              | HPN | DLB | JPN |
| 1            | 1                     | 0             | 0             | 2                | 3   | 3   | 2   |
| 2            | 0                     | 1             | 0             | 1                | 2   | 1   | 3   |
| 3            | 0                     | 0             | 1             | 4                | 4   | 5   | 6   |
| 4            | $\frac{1}{2}$         | $\frac{1}{2}$ | 0             | 3                | 1   | 2   | 1   |
| 5            | $\frac{1}{2}$         | 0             | $\frac{1}{2}$ | 5                | 5   | 4   | 4   |
| 6            | 0                     | $\frac{1}{2}$ | $\frac{1}{2}$ | 6                | 7   | 8   | 5   |
| 7            | $\frac{2}{3}$         | $\frac{1}{6}$ | $\frac{1}{6}$ | 7                | 6   | 9   | 8   |
| 8            | $\frac{1}{6}$         | $\frac{2}{3}$ | $\frac{1}{6}$ | 8                | 8   | 6   | 7   |
| 9            | $\frac{1}{6}$         | $\frac{1}{6}$ | $\frac{2}{3}$ | 10               | 9   | 10  | 9   |
| 10           | $\frac{1}{3}$         | $\frac{1}{3}$ | $\frac{1}{3}$ | 9                | 10  | 7   | 10  |

- (c) Construct and interpret a response trace plot for this experiment.  
 (d) Construct a response surface contour plot for the average rank response. What blends of the three varieties of grapes would you recommend?

**12.6** Reconsider the sweetener experiment in Exercise 12.2. Suppose that the experimenters are interested in a second response, the cost of the sweetener. For each of the 15 formulations of the product tested in the simplex-centroid design, the cost of the mixture of ingredients used in the formula and a manufacturing cost estimate are combined. For the design points, listed in the same order as in the original table in Exercise 12.2, these costs are as follows:

| Design Point | Cost | Design Point | Cost | Design Point | Cost |
|--------------|------|--------------|------|--------------|------|
| 1            | 4    | 6            | 8    | 11           | 12   |
| 2            | 15   | 7            | 16   | 12           | 18   |
| 3            | 6    | 8            | 13   | 13           | 15   |
| 4            | 25   | 9            | 18   | 14           | 16   |
| 5            | 10   | 10           | 15   | 15           | 15   |

- (a) Build an appropriate mixture model for the cost response.  
 (b) Construct a response trace plot for the cost response. Interpret this graph.  
 (c) Construct a response surface contour plot of the cost response. What practical advice could you give a decision maker based on this graph?  
 (d) Consider both the response surface for cost and the aftertaste response surface developed in Exercise 12.2. Suppose that your objective is to select a blend of ingredients so that the sweetener has low aftertaste while keeping the cost of the product below 12. What blend of sweeteners would you recommend?
- 12.7 Continuation of Exercise 12.6** Reconsider the cost response surface from Exercise 12.6. Suppose that your objective was to minimize the cost of the sweetener while keeping the aftertaste below 13. What formulation of ingredients would you recommend for this product?
- 12.8** Consider the second-order polynomial in two variables:

$$E(y) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{12}x_1x_2 + \beta_{11}x_1^2 + \beta_{22}x_2^2$$

- (a) Show that if the term  $\beta_0$  is replaced by  $\beta_0(x_1 + x_2)$ , if  $x_1^2$  is replaced by  $x_1(1 - x_2)$ , and if  $x_2^2$  is replaced by  $x_2(1 - x_1)$ , the canonical form of the quadratic mixture model results.  
 (b) Why are the substitutions described in (a) required in a mixture experiment?
- 12.9** When observations are collected only at the points of a  $\{q, m\}$  simplex lattice, the estimates of the parameters in the canonical polynomial are simple functions of the observed responses. Consider the  $\{3, 2\}$  lattice, let  $\bar{y}_i$  be the average of  $n_i$  observations at each vertex  $x_i = 1, x_j = 0$  for  $j \neq i$ , and let  $\bar{y}_{ij}$  be the average of  $n_{ij}$  observations at the edge midpoints or 50:50 binary blends  $x_i = x_j = \frac{1}{2}, x_k = 0$  for  $i < j < k = 1, 2, 3$ .

- (a) Show that on substituting  $\bar{y}_i$  and  $\bar{y}_{ij}$  for the expected response at each of the six design points, the following equations result:

$$\begin{aligned}\bar{y}_1 &= \beta_1, & \bar{y}_2 &= \beta_2, & \bar{y}_3 &= \beta_3 \\ \bar{y}_{12} &= \frac{1}{2}\beta_1 + \frac{1}{2}\beta_2 + \frac{1}{4}\beta_{12} \\ \bar{y}_{13} &= \frac{1}{2}\beta_1 + \frac{1}{2}\beta_3 + \frac{1}{4}\beta_{13} \\ \bar{y}_{23} &= \frac{1}{2}\beta_2 + \frac{1}{2}\beta_3 + \frac{1}{4}\beta_{23}\end{aligned}$$

- (b) Show that the solution to the equations given in part (a) are as follows:

$$\begin{aligned}b_1 &= \bar{y}_1, & b_2 &= \bar{y}_2, & b_3 &= \bar{y}_3 \\ b_{12} &= 4\bar{y}_{12} - 2(\bar{y}_1 + \bar{y}_2) \\ b_{13} &= 4\bar{y}_{13} - 2(\bar{y}_1 + \bar{y}_3) \\ b_{23} &= 4\bar{y}_{23} - 2(\bar{y}_2 + \bar{y}_3)\end{aligned}$$

Notice that the constants 4 and 2 in the last three equations do not depend on the number of replicates at each design point, but arise instead from the  $x_i = x_j = \frac{1}{2}$  values at each edge midpoint.

- (c) Use the equations derived in part (b) above to calculate the parameter estimates for the yarn elongation data in Example 12.1. Verify that the results obtained agree with the least squares estimates of these parameters obtained in Example 12.1.

**12.10** Consider the quadratic mixture model

$$y = \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{12}x_1x_2 + \beta_{13}x_1x_3 + \beta_{23}x_2x_3 + \varepsilon$$

where  $\varepsilon$  is the usual  $NID(0, \sigma^2)$  random error term. Suppose that the  $\{3, 2\}$  lattice design as described in Exercise 12.8 above is run, and the resulting data are used to estimate the coefficients in the quadratic model.

- (a) Show that

$$E(b_i) = \beta_i$$

and that

$$E(b_{ij}) = \beta_{ij}$$

- (b) Show that

$$\text{Var}(b_i) = \frac{\sigma^2}{n_i}, \quad i = 1, 2, 3$$

and

$$\text{Var}(b_{ij}) = \frac{16\sigma^2}{n_{ij}} + \frac{4\sigma^2}{n_i} + \frac{4\sigma^2}{n_j}, \quad i < j = 2, 3$$

(c) Show that

$$\text{Cov}(b_i, b_j) = 0, \quad i \neq j$$

$$\text{Cov}(b_i, b_{ij}) = \frac{-2\sigma^2}{n_i}$$

$$\text{Cov}(b_{ij}, b_{ik}) = \frac{4\sigma^2}{n_i}, \quad j \neq k$$

**12.11** Consider the simplex-centroid design for three components. Let  $\bar{y}_i$ ,  $\bar{y}_{ij}$ , and  $\bar{y}_{123}$  be the averages of  $n_i$  observations at the three vertices,  $n_{ij}$  observations at the edge midpoints, and  $n_{123}$  observations at the design centroid. The model of interest is

$$E(y) = \sum_{i=1}^3 \beta_i x_i + \sum_{i<j=2}^3 \sum_{i=1}^3 \beta_{ij} x_i x_j + \beta_{123} x_1 x_2 x_3$$

(a) Show that by substituting  $\bar{y}_i$ ,  $\bar{y}_{ij}$ , and  $\bar{y}_{123}$  for the expected response at each of the seven design points, the following equations result:

$$\bar{y}_i = \beta_i, \quad i = 1, 2, 3$$

$$\bar{y}_{ij} = \frac{1}{2}(\beta_i + \beta_j) + \frac{1}{4}\beta_{ij}, \quad i < j = 2, 3$$

$$\bar{y}_{123} = \frac{1}{3}(\beta_1 + \beta_2 + \beta_3) + \frac{1}{9}(\beta_{12} + \beta_{13} + \beta_{23}) + \frac{1}{27}\beta_{123}$$

(b) Show that the solution to the equation given in part (a) are

$$b_i = \bar{y}_i, \quad i = 1, 2, 3$$

$$b_{ij} = 4\bar{y}_{ij} - 2(\bar{y}_i + \bar{y}_j), \quad i < j = 2, 3$$

$$b_{123} = 27\bar{y}_{123} - 12(\bar{y}_{12} + \bar{y}_{13} + \bar{y}_{23}) + 9(\bar{y}_1 + \bar{y}_2 + \bar{y}_3)$$

(c) What effect does the observation at the centroid have on the estimates of the linear and quadratic terms in this model?

**12.12** Consider the special cubic model

$$\hat{y} = \sum_{i=1}^3 b_i x_i + \sum_{i<j=2}^3 \sum_{i=1}^3 b_{ij} x_i x_j + b_{123} x_1 x_2 x_3$$



- (a) How large must the  $b_{ij}$  term be if the term  $b_{12} x_i x_j$  is to contribute as much to the model at the edge midpoint as the linear term  $b_i x_i$ ?
- (b) How large must the cubic term  $b_{123}$  be if the contribution of the term  $b_{123} x_1 x_2 x_3$  to the model at the centroid is as large as the contribution of the quadratic term  $b_{ij} x_i x_j$ ?

**12.13** Consider the quadratic model in two components:

$$E(y) = \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

- (a) Suppose that four observations are collected at the following points:

$$y_1 \text{ at } x_1 = 1, x_2 = 0$$

$$y_2 \text{ at } x_1 = 0, x_2 = 1$$

$$y_3 \text{ and } y_4 \text{ at } x_1 = x_2 = \frac{1}{2}$$

Derive the estimates of the model coefficients for this design.

- (b) Suppose that four observations are collected at the following points:

$$y_1 \text{ at } x_1 = 1, x_2 = 0$$

$$y_2 \text{ at } x_1 = 0, x_2 = 1$$

$$y_3 \text{ at } x_1 = \frac{1}{3}, x_2 = \frac{2}{3}$$

$$y_4 \text{ at } x_1 = \frac{2}{3}, x_2 = \frac{1}{3}$$

Derive the estimates of the model coefficients for this design.

- (c) Which of these two designs would you prefer, if your objective was to find minimum variance estimates of the coefficients in the quadratic model?

**12.14** Reconsider the sweetener formulation problem in Exercise 12.2. Reparameterize the problem as was discussed in Section 12.4. Fit both linear and quadratic models, and verify that they can be converted to the Scheffé models you fitted in Exercise 12.2.

**12.15** Cornell (2000) presents an example of a three-component mixture experiment used to study the tint strength of a house paint as a function of the relative proportions of two pigments ( $P_1$  and  $P_2$ ) and a solvent ( $S$ ). The volume percentages of the three ingredients are fixed at 8%, so we can write the usual mixture constraint  $x_1 + x_2 + x_3 = 1$  by letting  $x_1 = P_1/8\%$ ,  $x_2 = P_2/8\%$ , and  $x_3 = S/8\%$ , respectively. This scaling produces what are often called **real** mixture components. The data given by Cornell are in Table E12.5.

Fit a second-order mixture model to the data, and find a blend that maximizes the tint strength.

**12.16** Reconsider the paint experiment in Exercise 12.15. Fit a second-order slack variable model to the data, using  $x_1$  as the slack variable. Verify that the contour plots from this model are identical to the ones obtained from the original Scheffé polynomial

TABLE E12.5 Data for Exercise 12.15

| Blend | $x_1 (P_1)$   | $x_2 (P_2)$   | $x_3 (S)$     | Tint Strength         |
|-------|---------------|---------------|---------------|-----------------------|
| 1     | 1             | 0             | 0             | 1.19                  |
| 2     | 0             | 1             | 0             | 0.97                  |
| 3     | 0             | 0             | 1             | -5.68                 |
| 4     | $\frac{1}{2}$ | $\frac{1}{2}$ | 0             | 0.13                  |
| 5     | $\frac{1}{2}$ | 0             | $\frac{1}{2}$ | -4.02                 |
| 6     | 0             | $\frac{1}{2}$ | $\frac{1}{2}$ | -4.18                 |
| 7     | $\frac{1}{3}$ | $\frac{1}{3}$ | $\frac{1}{3}$ | -3.36, -3.68<br>-3.04 |

in Exercise 12.15. Fit another second-order slack variable model, this time using  $x_2$  as the slack variable. Obtain the contour plots. Note that the model form is different, but the contour plots are identical to the ones from the original Scheffé polynomial.

- 12.17** Reconsider the paint experiment in Exercise 12.15. Fit a second-order slack variable model to the data using  $x_1$  as the slack variable. Eliminate the interaction term from this model, and refit. Construct a contour plot for this reduced model, and compare it with the contour plot for the full Scheffé model. Based on this analysis, how comfortable are you in general with the slack variable model approach?
- 12.18** Consider fitting a quadratic mixture model with three components. Suppose that you are considering two different designs, the  $\{3, 3\}$  lattice design and the augmented simplex lattice design in Fig. 12.12. Compare the two designs in terms of the precision of estimation of the quadratic mixture model terms. Which design would you prefer?
- 12.19** Consider fitting a quadratic mixture model with three components. Suppose that you are considering two different designs, the  $\{3, 3\}$  lattice design and the augmented simplex lattice design in Fig. 12.12. Compare the two designs in terms of the prediction variance of the response over the mixture space. Which design would you prefer?
- 12.20** Consider fitting a special cubic mixture model with three components. Suppose that you are considering two different designs, the  $\{3, 3\}$  lattice design and the augmented simplex lattice design in Fig. 12.12. Compare the two designs in terms of the prediction variance of the response over the mixture space. Which design would you prefer?
- 12.21** Consider the augmented simplex lattice design in Fig. 12.12. Suppose that the experimenter wants to modify this design by moving the location of the axial blends to 70% of the distance to the vertex instead of  $2/3$ , because this corresponds to blends that have some practical appeal to production personnel. What do you think of this modification? Compare the modified design to the original augmented simplex lattice design for a quadratic mixture model in terms of prediction variance over the mixture space.

- 12.22** Consider the augmented simplex lattice design in Exercise 12.21. Suppose that you could add four more runs to this design to obtain an estimate of pure error. One strategy would be to replicate the axial points and the centroid, a second alternative would be to replicate the vertices and the centroid, and a third strategy would be to replicate the edge centers and the centroid. Compare these strategies in terms of the standard errors of the model coefficients and the variance of the predicted response over the design region. Assume that the experimenter wants to fit a quadratic mixture model. What approach would you recommend?
- 12.23** Rework Exercise 12.21 assuming that the experimenter may need to fit a special cubic model.
- 12.24** Rework Exercise 12.22 assuming that the experimenter may need to fit a special cubic model.

### Short Answer Questions

- 12.25** In designing mixture experiments, a good rule of thumb is to assume that at least a second-order mixture model is likely to be adequate. **True False**
- 12.26** The simplex-lattice and simplex-centroid designs have most of the design points on the boundary of the experimental region. **True False**
- 12.27** When axial check blends are added to a simplex lattice, these blends are as far as possible (Euclidean distance) from the other design points. **True False**
- 12.28** The (3,2) simplex-lattice design is a rotatable design for fitting a second-order mixture model. **True False**
- 12.29** If one component has a response trace plot that is a horizontal line this implies that this component is inactive. **True False**
- 12.30** Scheffé mixture models do not contain an intercept term because this term would be perfectly correlated with the linear mixture terms. **True False**
- 12.31** A Scheffé mixture model can always be reparameterized to contain an intercept term. **True False**
- 12.32** The slack variable approach to fitting mixture models eliminates one of the mixture components from the model. **True False**



---

# 13

---

## OTHER MIXTURE DESIGN AND ANALYSIS TECHNIQUES

In Chapter 12 we focused on mixture experiments for exploring the complete simplex region. We introduced simplex-lattice and simplex-centroid designs that would enable us to fit response surface models over this region. In many mixture experiments there are **restrictions**, or **constraints**, on the component proportions that prevent us from exploring the entire simplex region. Frequently these restrictions take the form of lower and/or upper bounds on the component proportions. This problem is discussed in this chapter. In addition, we also briefly discuss several other variations of the mixture problem, including the use of **ratios of components** as the design variables, the inclusion of **process variables** in mixture experiments, and **screening techniques** to identify the most important components in the mixture.

### 13.1 CONSTRAINTS ON THE COMPONENT PROPORTIONS

In many mixture experiments there are constraints on the component proportions. These are often upper- and/or lower-bound constraints of the form  $L_i \leq x_i \leq U_i$ ,  $i = 1, 2, \dots, q$ , where  $L_i$  is the lower bound for the  $i$ th component and  $U_i$  is the upper bound for the  $i$ th component. Essentially,  $L_i$  represents a minimum proportion of the  $i$ th component, and  $U_i$  a maximum proportion of the  $i$ th component, that must be present in the mixture. These constraints are usually determined by practical bounds on the relative proportion of components that produce useable results. The general form of the constrained mixture problem is

$$\begin{aligned}x_1 + x_2 + \dots + x_q &= 1 \\L_i \leq x_i \leq U_i, & \quad i = 1, 2, \dots, q\end{aligned}\tag{13.1}$$

where  $L_i \geq 0$  and  $U_i \leq 1$  for  $i = 1, 2, \dots, q$ .

The effect of the upper- and lower-bound restriction in Equation 13.1 is to limit the feasible space for the mixture experiment to a subregion of the simplex. Therefore, experimental designs for constrained mixture problems must be limited to the feasible subregion of the simplex. In this section, we present design and analysis techniques for these constrained mixture spaces.

### 13.1.1 Lower-Bound Constraints on the Component Proportions

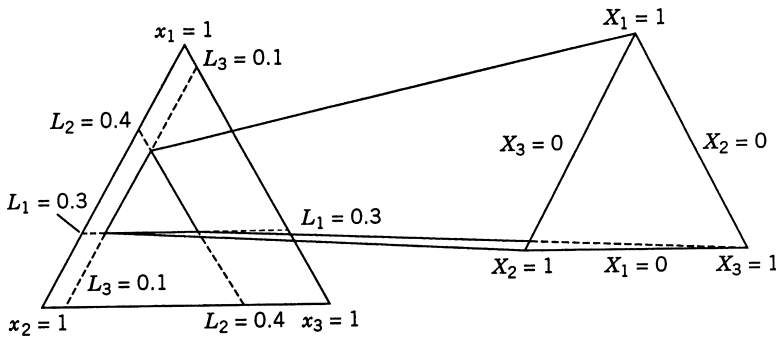
We now consider the case where only the lower bounds in Equation 13.1 are imposed, so that the constrained mixture problem becomes

$$\begin{aligned} x_1 + x_2 + \dots + x_q &= 1 \\ L_i \leq x_i \leq 1, \quad i &= 1, 2, \dots, q \end{aligned} \tag{13.2}$$

To illustrate the effect that lower bounds have on the feasible mixture design space, consider a three-component example with constraints

$$0.3 \leq x_1, \quad 0.4 \leq x_2, \quad \text{and} \quad 0.1 \leq x_3$$

Figure 13.1 shows the feasible mixture space after these constraints are imposed. Notice that the imposition of lower bounds does not affect the shape of the feasible mixture space; it is still a simplex. In general this will be the case; if only lower bounds are imposed on any of the component proportions, the feasible region for the mixture experiment will be a smaller simplex inscribed inside the original (or unconstrained) region. If all lower bounds are equal ( $L_1 = L_2 = \dots = L_q$ ), the centroid of the original simplex will also be the centroid of the smaller inscribed simplex.



**Figure 13.1** The feasible mixture space for three components with lower bounds, and the redefinition of the region as a simplex in the  $L$ -pseudocomponents. [Adapted from Cornell (2002), with the permission of the publisher.]

Because the feasible experimental region is still a simplex, it seems reasonable to define a new set of components that will take on the values 0 to 1 over the feasible region. This will make design construction and model fitting easier over the constrained region. These redefined components are called ***L*-pseudocomponents**, or sometimes just **pseudocomponents** (Crosier, 1984). The pseudocomponents  $X_i$  are defined using the following transformation:

$$X_i = \frac{x_i - L_i}{1 - L} \quad (13.3)$$

where

$$L = \sum_{i=1}^q L_i < 1$$

is the sum of all of the lower bounds. A value of  $X_i = 0$  now corresponds to the minimum allowable amount of the component  $i$ , while  $X_i = 1$  is the allowable maximum (obtained by setting all the other components at their minimum proportions). To illustrate, we may use the situation described previously and shown graphically in Fig. 13.1. The pseudocomponents are

$$X_1 = \frac{x_1 - 0.3}{0.2}, \quad X_2 = \frac{x_2 - 0.4}{0.2}, \quad X_3 = \frac{x_3 - 0.1}{0.2}$$

because  $L_1 = 0.3$ ,  $L_2 = 0.4$ ,  $L_3 = 0.1$ ,  $L = 0.3 + 0.4 + 0.1 = 0.8$ , and  $1 - L = 0.2$ . The factor space in terms of pseudocomponents is shown on the right-hand side of Fig. 13.1. The orientation of the  $L$ -pseudocomponent simplex is the same as the original simplex, and the height of the  $L$ -pseudocomponent simplex is  $1 - L$ .

Constructing a design in the  $L$ -pseudocomponents is easy. One simply specifies the design points in the  $X_i$  and then converts them to the settings in the original components using

$$x_i = L_i + (1 - L)X_i \quad (13.4)$$

To illustrate, suppose that we decide to use a simplex-centroid design for our experiment. Table 13.1 shows the design points in the pseudocomponents, along with the corresponding setting for the original components.

The original component settings in Table 13.1 were calculated by using Equation 13.4. This becomes

$$\begin{aligned} x_1 &= 0.3 + 0.2X_1 \\ x_2 &= 0.4 + 0.2X_2 \\ x_3 &= 0.1 + 0.2X_3 \end{aligned}$$

**TABLE 13.1 Pseudocomponent Settings and Original Component Settings, Simplex-Centroid Design**

| Pseudocomponents |               |               | Original Components |        |        |
|------------------|---------------|---------------|---------------------|--------|--------|
| $X_1$            | $X_2$         | $X_3$         | $x_1$               | $x_2$  | $x_3$  |
| 1                | 0             | 0             | 0.5                 | 0.4    | 0.1    |
| 0                | 1             | 0             | 0.3                 | 0.6    | 0.1    |
| 0                | 0             | 1             | 0.3                 | 0.4    | 0.3    |
| $\frac{1}{2}$    | $\frac{1}{2}$ | 0             | 0.4                 | 0.5    | 0.1    |
| $\frac{1}{2}$    | 0             | $\frac{1}{2}$ | 0.4                 | 0.4    | 0.2    |
| 0                | $\frac{1}{2}$ | $\frac{1}{2}$ | 0.3                 | 0.5    | 0.2    |
| $\frac{1}{3}$    | $\frac{1}{3}$ | $\frac{1}{3}$ | 0.3667              | 0.4667 | 0.1666 |

For example, consider the point  $X_1 = \frac{1}{2}$ ,  $X_2 = \frac{1}{2}$ , and  $X_3 = 0$ . In terms of the original components, this point becomes

$$\begin{aligned}x_1 &= 0.3 + 0.2 \left( \frac{1}{2} \right) = 0.4 \\x_2 &= 0.4 + 0.2 \left( \frac{1}{2} \right) = 0.5 \\x_3 &= 0.1 + 0.2(0) = 0.1\end{aligned}$$

as shown in Table 13.1.

We recommend using pseudocomponents to fit the mixture model. The reason for this is that constrained design spaces usually have moderately high to high levels of **multicollinearity** or **ill-conditioning**, and this can have serious effects on the least squares estimators of regression coefficients in that the variances of these coefficient estimators are inflated. In general, a mixture model in the pseudocomponents will have lower levels of multicollinearity than will the same model in the original component proportions. For more discussion of this point, see Montgomery and Voth (1994) and St. John (1984). It is, of course, a relatively simple matter to convert a model in the pseudocomponents into the corresponding model in the actual components.

**Example 13.1 Formulating a Propellant** We use a three-component example from Montgomery and Voth (1994) to illustrate fitting and interpreting a model with pseudocomponents. The example includes blending fuel ( $x_1$ ), oxidizer ( $x_2$ ), and binder ( $x_3$ ) together to form a propellant used in aircrew escape systems. As in most mixture experiments, several responses are of interest, including both physical and mechanical properties of the propellant; however, we concentrate on three responses. The first is the burning rate ( $y_1$ , cm/sec), a critical characteristic if the escape system is to perform satisfactorily. It is measured by testing several samples of propellant from the same run and reporting the average observed burning rate. The second response is the standard deviation of the observed burning rates at each run ( $y_2$ , cm/sec); low variability is desirable, implying consistent performance across different escape systems assembled from propellant grains from the same batch. The third



response is a manufacturability index ( $y_3$ ) that reflects the cost and difficulty associated with producing a particular mixture. The constraints for this mixture problem are

$$\begin{aligned} x_1 + x_2 + x_3 &= 0.9 \\ 0.30 &\leq x_1 \\ 0.20 &\leq x_2 \\ 0.20 &\leq x_3 \end{aligned}$$

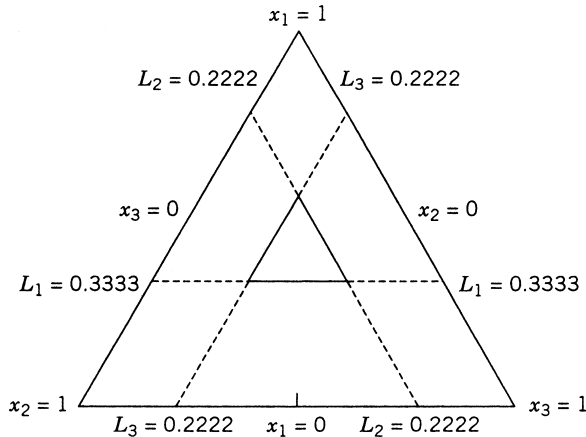
Notice that the three components must make up 90% of the mixture. The remaining 10% is a fourth ingredient that is never varied. We model this system in terms of pseudocomponents, say

$$X_i = \frac{x_i - L_i}{0.9 - \sum_{i=1}^3 L_i}, \quad i = 1, 2, 3 \tag{13.5}$$

This transformation yields pseudocomponents  $X_i$  such that  $0 \leq X_i \leq 1$ ,  $i = 1, 2, 3$ , and  $X_1 + X_2 + X_3 = 1$ . Because all the constraints in this problem are lower-bound constraints, the experimental region is a simplex. To draw this region in the usual coordinate system, we must remember that the range of each  $x_i$  on the graph is from 0 to 1. Therefore, the constraints that should be plotted on the graph are  $0.3/0.9 \leq x_1$ ,  $0.2/0.9 \leq x_2$ , and  $0.2/0.9 \leq x_3$  or

$$\begin{aligned} 0.3333 &\leq x_1 \\ 0.2222 &\leq x_2 \\ 0.2222 &\leq x_3 \end{aligned}$$

We can think of these constraints as the proportion of the flexible proportion of each component after the automatic inclusion of the mandatory 10%. This constrained region is shown in Fig. 13.2.



**Figure 13.2** The constrained design region for Example 13.1.

The experimenters decided that a quadratic mixture model would likely be adequate to describe the relationship between the responses and the three component proportions. They selected a 10-point design consisting of the simplex centroid augmented with the interior axial points located midway between the centroid and the opposed vertex. They decided to replicate each vertex twice and the centroid three times, resulting in the final 15-run design shown in Table 13.2. The relationship between the settings in the pseudocomponents and the actual component settings is

$$x_i = L_i + X_i \left( 0.9 - \sum_{i=1}^3 L_i \right)$$

For the first design point in Table 13.2, where  $X_1 = 1, X_2 = X_3 = 0$  we have  $x_1 = 0.5, x_2 = 0.2$ , and  $x_3 = 0.2$ .

Table 13.3 shows the results of sequentially fitting the linear, quadratic, and special cubic models to the data in Table 13.2, using the mean burning rate response  $y_1$ . Based on the lack-of-fit test and the small  $P$ -value associated with the special cubic term, we would recommend the special cubic model. Note also that the value of the PRESS statistic is smallest for the special cubic model. As noted in Chapter 2, regression models with small values of PRESS are usually good prediction equations.

In judging the relative size of PRESS, recall that it is often useful to compute the  $R^2$  for prediction based on PRESS:

$$R^2_{\text{prediction}} = 1 - \frac{\text{PRESS}}{SS_{\text{Total}}}$$

This statistic can be thought of as a measure of the capability of the model to predict new data, and it can be compared with the usual  $R^2$  and  $R^2$ -adjusted statistics. For the special cubic model we have

$$R^2_{\text{prediction}} = 1 - \frac{3296.975}{9557.778} = 0.6550$$

This is somewhat smaller than the other two  $R^2$ -values (see Table 13.3).

Table 13.4 contains a summary of the special cubic model. The fitted model in the pseudocomponents is

$$\begin{aligned} \hat{y}_1 = & 35.49X_1 + 42.78X_2 + 70.36X_3 + 16.02X_1X_2 \\ & + 36.33X_1X_3 + 136.82X_2X_3 + 854.98X_1X_2X_3 \end{aligned}$$

By substituting the definitions of the pseudocomponents into this equation, we may obtain a model for the mean burning rate in terms of the original (actual) components. However, it is usually not necessary to do this, because modern software will plot the response surface contours in either the pseudocomponents or the actual components.

From Table 13.4, notice that in the final model there are three values of the hat diagonals  $h_{ii}$  that are large (0.91); these are associated with the centers-of-edges design points. This has occurred because the  $h_{ii}$  are both model- and design-dependent, and the final model here is of higher order than the one that was initially contemplated. This happens frequently

**TABLE 13.2 Data for the Propellant Example**

| Observation | Pseudocomponents |        |        | Original Components |                  |                | Standard Deviation of |                    | Manufacturability Index $y_3$ |
|-------------|------------------|--------|--------|---------------------|------------------|----------------|-----------------------|--------------------|-------------------------------|
|             | $X_1$            | $X_2$  | $X_3$  | $x_1$ (Fuel)        | $x_2$ (Oxidizer) | $x_3$ (Binder) | Burning Rate $y_1$    | Burning Rate $y_2$ |                               |
| 1           | 1.0              | 0.0    | 0.0    | 0.5                 | 0.2              | 0.2            | 32.5                  | 4.1                | 32                            |
| 2           | 1.0              | 0.0    | 0.0    | 0.5                 | 0.2              | 0.2            | 37.9                  | 3.7                | 25                            |
| 3           | 0.5              | 0.5    | 0.0    | 0.4                 | 0.3              | 0.2            | 44.0                  | 6.8                | 20                            |
| 4           | 0.5              | 0.0    | 0.5    | 0.4                 | 0.2              | 0.3            | 63.2                  | 4.7                | 18                            |
| 5           | 0.0              | 1.0    | 0.0    | 0.3                 | 0.4              | 0.2            | 54.5                  | 8.9                | 18                            |
| 6           | 0.0              | 1.0    | 0.0    | 0.3                 | 0.4              | 0.2            | 32.5                  | 9.2                | 21                            |
| 7           | 0.0              | 0.5    | 0.5    | 0.3                 | 0.3              | 0.3            | 94.0                  | 4.5                | 17                            |
| 8           | 0.0              | 0.0    | 1.0    | 0.3                 | 0.2              | 0.4            | 64.0                  | 14.0               | 14                            |
| 9           | 0.0              | 0.0    | 1.0    | 0.3                 | 0.2              | 0.4            | 78.5                  | 13.0               | 16                            |
| 10          | 0.6667           | 0.1667 | 0.1667 | 0.4333              | 0.2333           | 0.2333         | 67.1                  | 3.5                | 20                            |
| 11          | 0.1667           | 0.6667 | 0.1667 | 0.3333              | 0.3333           | 0.2333         | 73.0                  | 5.2                | 22                            |
| 12          | 0.1667           | 0.1667 | 0.6667 | 0.3333              | 0.2333           | 0.3333         | 87.5                  | 7.0                | 17                            |
| 13          | 0.3333           | 0.3333 | 0.3333 | 0.3667              | 0.2667           | 0.2667         | 112.5                 | 4.6                | 19                            |
| 14          | 0.3333           | 0.3333 | 0.3333 | 0.3667              | 0.2667           | 0.2667         | 98.5                  | 3.5                | 20                            |
| 15          | 0.3333           | 0.3333 | 0.3333 | 0.3667              | 0.2667           | 0.2667         | 103.6                 | 3.0                | 18                            |

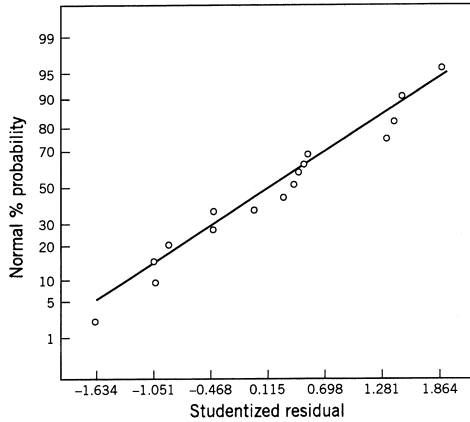
TABLE 13.3 Analysis of Variance Summary and Lack-of-Fit Tests for Mean Burning Rate Response

| <i>ANOVA Summary of Model Fit</i> |                |                    |             |         |                 |                |                         |            |  |
|-----------------------------------|----------------|--------------------|-------------|---------|-----------------|----------------|-------------------------|------------|--|
| Source                            | Sum of Squares | Degrees of Freedom | Mean Square | F-Value | Probability > F | R <sup>2</sup> | Adjusted R <sup>2</sup> | PRESS      |  |
| Intercept                         | 72,564.993     | 1                  | 72,564.993  |         |                 |                |                         |            |  |
| Linear                            | 2,395.909      | 2                  | 1,197.954   | 2.007   | 0.1770          | 0.2507         | 0.1258                  | 11,323.250 |  |
| Quadratic                         | 5,486.852      | 3                  | 1,828.951   | 9.827   | 0.0034          | 0.8247         | 0.7274                  | 8,284.901  |  |
| Special cubic                     | 1,063.615      | 1                  | 1,063.615   | 13.92   | 0.0058          | 0.9360         | 0.8881                  | 3,296.975  |  |
| Error                             | 611.402        | 8                  | 76.425      |         |                 |                |                         |            |  |
| Corrected total                   | 9,557.778      | 14                 |             |         |                 |                |                         |            |  |
| <i>Lack-of-Fit Tests</i>          |                |                    |             |         |                 |                |                         |            |  |
| Model                             | Sum of Squares | DF                 | Mean Square | F-Value | Probability > F |                |                         |            |  |
| Linear                            | 6,699.757      | 7                  | 957.108     | 10.36   | 0.0102          |                |                         |            |  |
| Quadratic                         | 1,212.905      | 4                  | 303.226     | 3.281   | 0.1123          |                |                         |            |  |
| Special cubic                     | 149.290        | 3                  | 49.763      | 0.5384  | 0.6763          |                |                         |            |  |
| Pure error                        | 462.112        | 5                  | 92.422      |         |                 |                |                         |            |  |

**TABLE 13.4 Results of Fitting the Special Cubic Model for the Mean Burning Rate Response**

| Variable    | Parameter Estimate | Degrees of Freedom | Standard Error | $t$ for $H_0$ Parameter = 0 | Probability $> t $ | Observation | Actual Value | Predicted Value | Residual  | $h_{ii}$ | Studentized Residual | Cook's Distance | PRESS Residual |
|-------------|--------------------|--------------------|----------------|-----------------------------|--------------------|-------------|--------------|-----------------|-----------|----------|----------------------|-----------------|----------------|
| $x_1$       | 35.49456           | 1                  | 6.07214        | 5.845                       | 0.0004             | 1           | 32.50000     | 35.49456        | -2.99456  | 0.482    | -0.476               | 0.030           | -5.78100       |
| $x_2$       | 42.77552           | 1                  | 6.07214        | 7.045                       | 0.0001             | 2           | 37.90000     | 35.49456        | 2.40544   | 0.482    | 0.382                | 0.019           | 4.64371        |
| $x_3$       | 70.36123           | 1                  | 6.07214        | 11.59                       | 0.0001             | 3           | 44.00000     | 43.14016        | 0.85984   | 0.910    | 0.327                | 0.154           | 9.55378        |
| $x_1x_2$    | 16.02049           | 1                  | 38.29236       | 0.4184                      | 0.6867             | 4           | 63.20000     | 62.01159        | 1.18841   | 0.910    | 0.452                | 0.294           | 13.2046        |
| $x_1x_3$    | 36.33478           | 1                  | 38.29236       | 0.9489                      | 0.3705             | 5           | 54.50000     | 42.77552        | 11.72448  | 0.482    | 1.864                | 0.463           | 22.63413       |
| $x_2x_3$    | 136.82049          | 1                  | 38.29236       | 3.573                       | 0.0073             | 6           | 32.50000     | 42.77552        | -10.27552 | 0.482    | -1.634               | 0.355           | -19.83691      |
| $x_1x_2x_3$ | 854.98182          | 1                  | 229.18322      | 3.731                       | 0.0058             | 7           | 94.00000     | 90.77350        | 3.22650   | 0.910    | 1.227                | 2.165           | 35.85006       |
|             |                    |                    |                |                             |                    | 8           | 64.00000     | 70.36123        | -6.36123  | 0.482    | -1.011               | 0.136           | -12.28037      |
|             |                    |                    |                |                             |                    | 9           | 78.50000     | 70.36123        | 8.13877   | 0.482    | 0.886                | 0.223           | 15.71191       |
|             |                    |                    |                |                             |                    | 10          | 67.10000     | 67.96999        | -0.86999  | 0.186    | -0.110               | 0.000           | -1.06878       |
|             |                    |                    |                |                             |                    | 11          | 73.00000     | 79.98427        | -6.98427  | 0.186    | -0.886               | 0.026           | -8.56018       |
|             |                    |                    |                |                             |                    | 12          | 87.50000     | 95.46999        | -7.96999  | 0.786    | -1.011               | 0.033           | -9.79114       |
|             |                    |                    |                |                             |                    | 13          | 112.50000    | 102.22929       | 10.27071  | 0.273    | 1.378                | 0.102           | 14.12752       |
|             |                    |                    |                |                             |                    | 14          | 98.50000     | 102.22929       | -3.72929  | 0.273    | -0.500               | 0.013           | -5.12970       |
|             |                    |                    |                |                             |                    | 15          | 103.60000    | 102.22929       | 1.37071   | 0.273    | 0.184                | 0.002           | 1.88543        |

Standard error of mean = 2.25721.

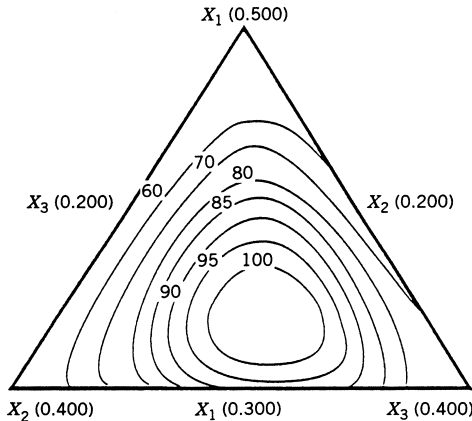


**Figure 13.3** Normal probability plot of the studentized residuals for the mean burning rate response.

in mixture experiments and often has undesirable consequences. For example, the large  $h_{ii}$  value for design point 7 combined with a moderately large studentized residual, impacts the Cook's distance measure and the PRESS residual. The large PRESS residual dramatically reduces the  $R^2$  prediction. Hence this single observation exerts disproportionate influence on the model and its estimates.

Figure 13.3 presents a normal probability plot of the studentized residuals from the special cubic model. We recommend plotting studentized residuals from mixture models, because the highly nonuniform distribution of leverage in many mixture designs results in residuals that may have very different variances. Thus, in a normal probability plot, the studentized residuals should lie approximately along a straight line. Figure 13.4 presents a contour plot for the mean burning rate response.

Table 13.5 summarizes the model-building activity for the other two responses: standard deviation of burning rate ( $y_2$ ) and manufacturability index ( $y_3$ ). In both cases, we selected the model with minimum PRESS, resulting in a quadratic model for the standard deviation



**Figure 13.4** Mean burning rate: contour plot.

**TABLE 13.5 Mixture Models for Standard Deviation of Burning Rate ( $y_2$ ) and Manufacturability Index ( $y_3$ )**

| Model Form    | $y_2$  |        | $y_3$  |        |
|---------------|--------|--------|--------|--------|
|               | $R^2$  | PRESS  | $R^2$  | PRESS  |
| Linear        | 0.4686 | 150.60 | 0.7942 | 39.83  |
| Quadratic     | 0.9830 | 7.15   | 0.8222 | 69.50  |
| Special cubic | 0.9844 | 13.47  | 0.8337 | 140.51 |

$$\hat{y}_2 = 3.88159X_1 + 9.03873X_2 + 13.63397X_3 - 0.19048X_1X_2 - 16.61905X_1X_3 - 27.67619X_2X_3$$

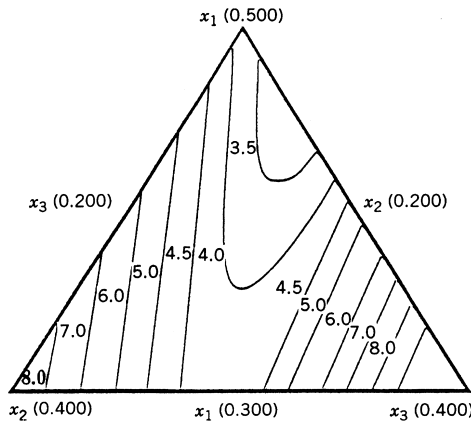
$$R^2_{\text{prediction}} = 0.9575$$

$$\hat{y}_3 = 23.13333X_1 + 19.73333X_2 + 14.73333X_3$$

$$R^2_{\text{prediction}} = 0.6456$$

of burning rate and a linear model for the manufacturability index. The values of  $R^2_{\text{prediction}}$  for those models, also shown in Table 13.5, are reasonably close to the ordinary  $R^2$ -statistic for the selected model. Plots of the studentized residuals for both models were satisfactory, and the contour plots are shown in Figs. 13.5 and 13.6.

As in most mixture designs, here the experimenters are interested in formulating a product that satisfies several customer requirements. Specifically, the mean burning rate must exceed 95 cm/sec, the standard deviation of the burning rate should be small (less than 4.5 cm/sec, say), and the manufacturability index should not exceed 20. Figure 13.7 presents a plot of the design region with these constraints superimposed. This graph identifies the feasible mixtures that satisfy all three response constraints. Obviously, there are many product formulations that will be satisfactory. The final choice between formulations could now be made using other supplemental criteria. For example, the conclusions here are based on predicted mean responses, and no measure of uncertainty in these predictions was evaluated. One could use confidence or prediction intervals to represent the uncertainty. Alternatively, the effects of tightening the constraints can be investigated. Figure 13.8 illustrates the effects of requiring the standard deviation of the burning rate to be less than



**Figure 13.5** Standard deviation of burning rate: contour plot.

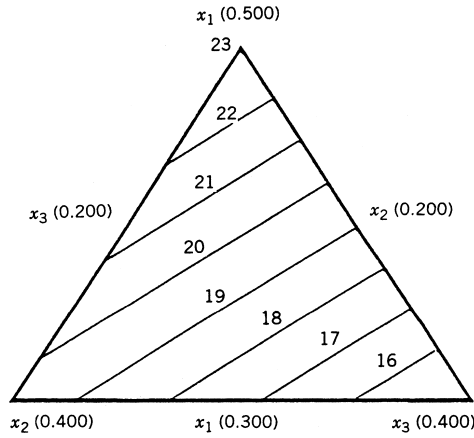


Figure 13.6 Manufacturability index: contour plot.

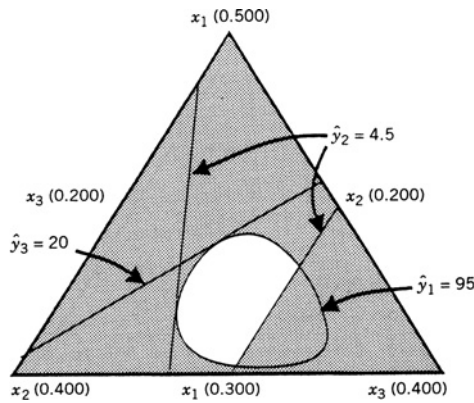


Figure 13.7 Region of optimum solution for the propellant formulation problem.

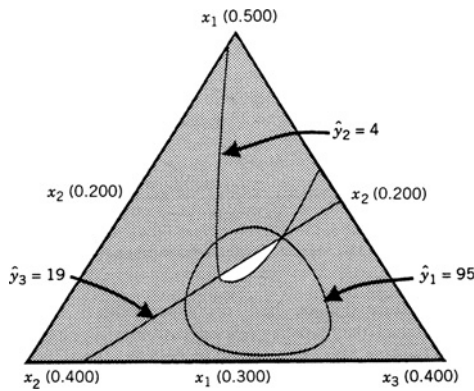


Figure 13.8 Region of optimum solution with tighter constraints on standard deviation of burning rate and manufacturability index.



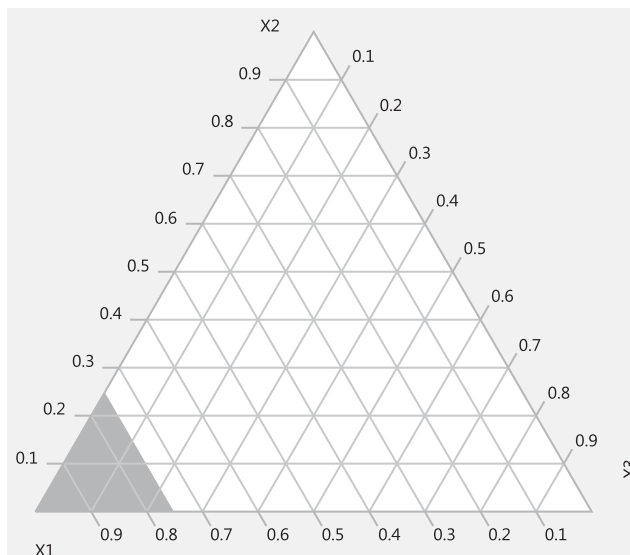
4 cm/sec and the manufacturability index to be less than 19. This new set of constraints has greatly reduced the feasible mixture space, and it is approaching the practical limits of reduction in both variability and cost for this product if the burning rate requirement is to be met. Alternately, one could use some of the multiple response strategies in Chapter 7.

Finally, the effect of the influential point (design point 7) may be investigated. A simple way to do this is to withhold the point, refit the model (or models) in question, and examine the changes that occur. When this analysis is performed, similar results are obtained, but the feasible region from this refitting is somewhat smaller. It would be prudent to select a final formulation that falls in the feasible region from both models as this offers some protection from the overinfluence of design point 7.

### 13.1.2 Upper-Bound Constraints on the Component Proportions

Sometimes only **upper-bound constraints** of the form  $x_i \leq U_i$  are placed on the component proportions. In these cases, a simple modification of the simplex-lattice design consists in replacing the restricted components with mixtures consisting of combinations of the restricted components and predetermined combinations of the unrestricted components. This approach is discussed and illustrated in Cornell (2002). However, as is obvious from inspection of Fig. 13.9, where a single upper-bound constraint  $x_1 \leq 0.75$  has been imposed, the presence of upper bounds can lead to an experimental region that is not a simplex (notice that the JMP ternary plot has the axes labeled differently than we have done previously). In such cases, **computer-generated designs** (typically based on either the  $D$ - or  $I$ -optimality criteria) are logical design alternatives.

Table 13.6 is a 12-run  $D$ -optimal design (found using JMP) assuming a quadratic mixture model for the mixture experiment in Fig. 13.9 having only an upper bound constraint. The design is shown graphically in Fig. 13.10. Notice that while there are 12 runs, there are only nine distinct design points. Runs 3 and 11, 6 and 7, and 8 and 9 are replicates. Two

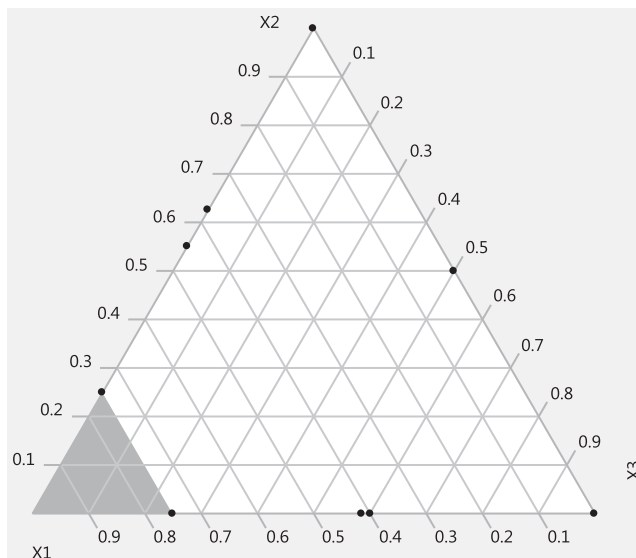


**Figure 13.9** A three-component mixture experiment with an upper bound constraint  $x_1 \leq 0.75$ .

**TABLE 13.6** A 12-Run *D*-Optimal Design for the Mixture Experiment in Fig. 13.9

| Run | $x_1$        | $x_2$ | $x_3$        |
|-----|--------------|-------|--------------|
| 1   | 0.75         | 0     | 0.25         |
| 2   | 0.75         | 0.25  | 0            |
| 3   | 0            | 0.5   | 0.5          |
| 4   | 0.4142829008 | 0     | 0.5857170992 |
| 5   | 0.398304425  | 0     | 0.601695575  |
| 6   | 0            | 0     | 1            |
| 7   | 0            | 0     | 1            |
| 8   | 0            | 1     | 0            |
| 9   | 0            | 1     | 0            |
| 10  | 0.375        | 0.625 | 0            |
| 11  | 0            | 0.5   | 0.5          |
| 12  | 0.45         | 0.55  | 0            |

of these pairs of runs are vertices and the third pair is at the center of the flat connecting the  $x_2$  and  $x_3$  vertices. There are no runs in the interior of the simplex. Table 13.7 is the 12-run *I*-optimal design for this experiment. The design is shown graphically in Fig. 13.11. There are 11 distinct design points. Only runs 6 and 10 are replicates. As in the *D*-optimal design, these are the runs at the center of the flat connecting the  $x_2$  and  $x_3$  vertices. Notice that in contrast to the *D*-optimal design, the *I*-optimal design has two runs in the interior of the simplex. Figure 13.12 is a fraction of design space plot for these two designs. As anticipated, the *I*-optimal design has considerably smaller prediction variance over most of the design space in comparison to the *D*-optimal design. The *D*-optimal design only has smaller prediction variance over the last 5% or so of the design space.

**Figure 13.10** A Ternary Plot of the 12-Run *D*-Optimal Design for the Mixture Experiment in Table 13.6.

**TABLE 13.7 A 12-Run I-Optimal Design for the Mixture Experiment in Fig. 13.9**

| Run | $x_1$        | $x_2$    | $x_3$        |
|-----|--------------|----------|--------------|
| 1   | 0.26025      | 0.369875 | 0.369875     |
| 2   | 0.375        | 0.625    | 0            |
| 3   | 0            | 1        | 0            |
| 4   | 0.406        | 0.594    | 0            |
| 5   | 0.3898557404 | 0        | 0.6101442596 |
| 6   | 0            | 0.5      | 0.5          |
| 7   | 0.275        | 0.3625   | 0.3625       |
| 8   | 0            | 0        | 1            |
| 9   | 0.75         | 0.25     | 0            |
| 10  | 0            | 0.5      | 0.5          |
| 11  | 0.4231211765 | 0        | 0.5768788235 |
| 12  | 0.75         | 0        | 0.25         |

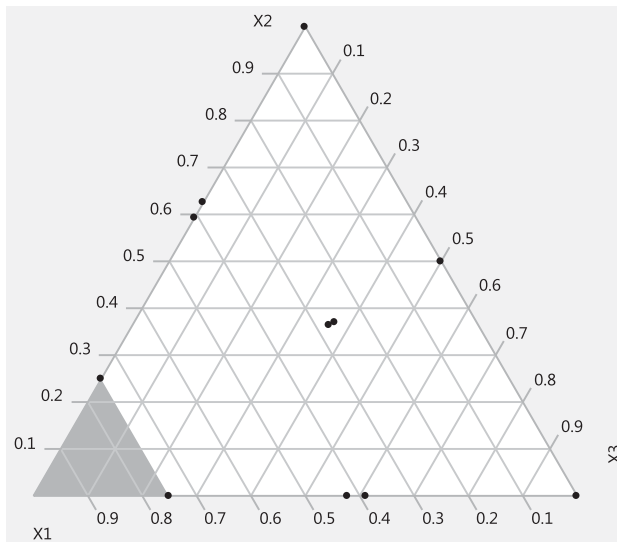
In some cases, upper-bound constraints lead to a design space that is a simplex. For example, consider the three-component problem

$$x_1 \leq 0.4, \quad x_2 \leq 0.5, \quad \text{and} \quad x_3 \leq 0.3$$

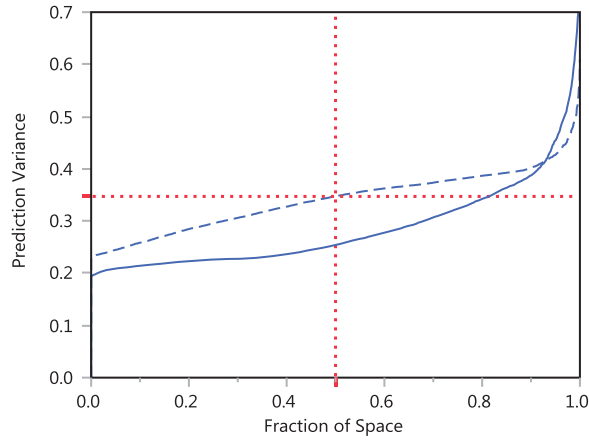
is shown in Fig. 13.13a. Notice that the feasible region for this mixture experiment is an inverted simplex. Obviously a standard simplex-type design can be defined over this region Fig. 13.13b shows the feasible region for the three-component problem:

$$x_1 \leq 0.7, \quad x_2 \leq 0.5, \quad \text{and} \quad x_3 \leq 0.8$$

In this case, the feasible experimental region is not a simplex, and, once again, a computer-generated design would be a logical choice.



**Figure 13.11** A ternary plot of the 12-run *I*-optimal design for the mixture experiment in Table 13.7.



**Figure 13.12** Fraction of design space plot for the mixture designs in Tables 13.6 and 13.7. The *D*-optimal design is the dashed line and the *I*-optimal design is the solid line.

In general, when upper-bound constraints  $x_1 \leq U_1$  are present, the feasible region will be an inverted simplex that lies entirely within the original or unconstrained simplex if and only if

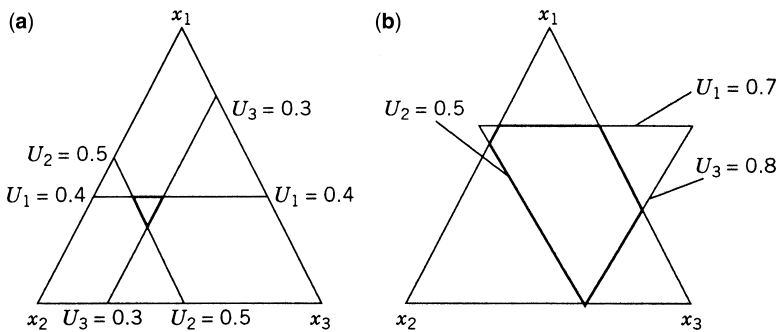
$$\sum_{i=1}^q U_i - U_{\min} \leq 1 \tag{13.6}$$

where  $U_{\min}$  is the smallest of the upper bounds  $U_i$ . For example, for the situation in Fig. 13.13a we have

$$\sum_{i=1}^3 U_i = 0.4 + 0.5 + 0.3 = 1.2$$

and  $U_{\min} = 0.3$ , so

$$\sum_{i=1}^3 U_i - U_{\min} = 1.2 - 0.3 = 0.9 < 1$$



**Figure 13.13** Three-component mixture experiments with upper-bound constraints. (a) The feasible region is an inverted simplex. (b) The feasible region is not a simplex.

and the feasible region is an inverted simplex. On the other hand, in Fig. 13.13b we have

$$\sum_{i=1}^3 U_i = 0.7 + 0.5 + 0.8 = 2.0$$

and  $U_{\min} = 0.5$ , so

$$\sum_{i=1}^3 U_i - U_{\min} = 2.0 - 0.5 = 1.5 > 1$$

and the feasible region is not a simplex.

When two or more of the component proportions are constrained by upper bounds, Crosier (1984) suggests defining upper-bound pseudocomponents, or  $U$ -pseudocomponents:

$$u_i = \frac{U_i - x_i}{\sum_{i=1}^q U_i - 1}, \quad i = 1, 2, \dots, q \quad (13.7)$$

where

$$\sum_{i=1}^q U_i > 1$$

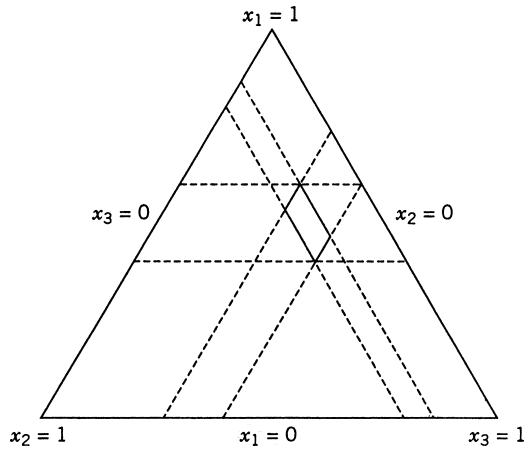
The region of the  $U$ -pseudocomponents is an inverted simplex; and when this inverted simplex lies entirely within the original simplex, it is easy to set up a standard simplex-type design in the  $U$ -pseudocomponents. The relationship between the settings in the original component proportions and the design points in the  $U$ -pseudocomponents  $U_i$  is

$$x_i = U_i - \left( \sum_{i=1}^q U_i - 1 \right) u_i \quad (13.8)$$

One then typically fits a model in the  $U$ -pseudocomponents. However, because the orientation of this design space is inverted, one must be careful in interpreting the coefficients in the fitted model when making inferences about the fitted surface in the original components.

### 13.1.3 Active Upper- and Lower-Bound Constraints

We now consider the case where there are both lower and upper bounds on the component proportions. In these cases, the feasible mixture region is no longer a simplex. To illustrate,



**Figure 13.14** Feasible experimental region for the shampoo foam experiment.

suppose that we wish to formulate a shampoo in terms of three component proportions  $x_1$  = lauryl sulfate,  $x_2$  = cocamide, and  $x_3$  = lauramide, and such that

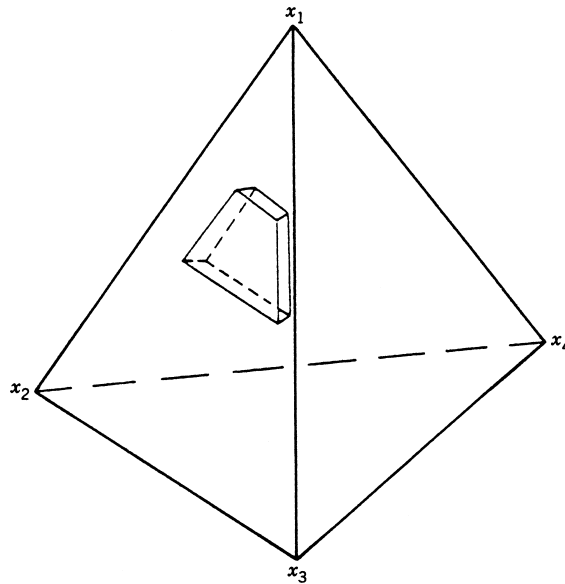
$$\begin{aligned}x_1 + x_2 + x_3 &= 0.50 \\0.20 &\leq x_1 \leq 0.30 \\0.07 &\leq x_2 \leq 0.10 \\0.13 &\leq x_3 \leq 0.20\end{aligned}$$

The remaining 50% of the mixture consists of water, perfume, and coloring agents whose proportions will be held fixed. The response variable of interest is the foam height of the shampoo.

Figure 13.14 shows the feasible region for this shampoo foam experiment. Remember that in constructing this graph, the components must be scaled so that  $x_1 + x_2 + x_3 = 1$ , so the upper and lower bounds on this graph were obtained by dividing the original upper and lower bounds by 0.5. Notice that the region is not a simplex; therefore, standard simplex-type mixture designs cannot be used.

As another example, consider the flare experiment described by McLean and Anderson (1966). The objective of the experiment was to determine the formulation of a railroad flare so as to maximize the illumination level. The flare consists of four ingredients:  $x_1$  = magnesium,  $x_2$  = sodium nitrate,  $x_3$  = strontium nitrate, and  $x_4$  = binder. The constraints on the four component proportions were

$$\begin{aligned}x_1 + x_2 + x_3 + x_4 &= 1 \\0.40 &\leq x_1 \leq 0.60 \\0.10 &\leq x_2 \leq 0.50 \\0.10 &\leq x_3 \leq 0.50 \\0.03 &\leq x_4 \leq 0.08\end{aligned}$$



**Figure 13.15** Constrained experimental region for the flare experiment. [Adapted from Cornell (2002), with permission of the publisher.]

The constrained experimental region is shown in Fig. 13.14. Notice that the feasible region for this experiment is not a simplex, and that the actual upper bound for  $x_2$  and  $x_3$  is 0.47, since any value larger than this would combine with the other minimum values for the other components to violate the components summing to one.

In general, when both upper- and lower-bound constraints are active in a mixture experiment, the feasible design space is an irregular hyperpolytope, and a computer-generated design should be used in the experiment. Generally either an  $I$ -optimal or a  $D$ -optimal design would be appropriate. Anderson and McLean suggest using an extreme vertices design, which we now briefly describe.

The extreme vertices of the constrained region are formed by the combinations of the upper- and lower-bound constraints. In the shampoo foam study ( $q = 3$  components) there are four extreme vertices shown in Fig. 13.14, and in the flare problem ( $q = 4$  components) there are eight extreme vertices shown in Fig. 13.15. McLean and Anderson (1966) proposed using these points as the basis of a design along with a subset of the remaining centroids or points that are at the centers of the edges, faces, and so forth, including the overall centroid of the region.

McLean and Anderson (1966) demonstrated this strategy with a 15-run design for the flare problem, assuming that a quadratic model would be adequate to describe the response. JMP has implemented the extreme vertices strategy and the 15-run JMP design is shown in Table 13.8.

While the extreme vertices strategy may generate an intuitively reasonable design (Snee and Marquardt, 1974), we believe that  $I$ -optimal or  $D$ -optimal designs are superior. To illustrate, Table 13.9 is a 15-run  $I$ -optimal design for the flare problem assuming a quadratic model. Figure 13.16 is the fraction of design space plot comparing the prediction variance performance of these two designs. Clearly, the  $I$ -optimal design is superior to the extreme vertices design in this situation.

**TABLE 13.8 A 15-Run Extreme Vertices Design for the Flair Problem**

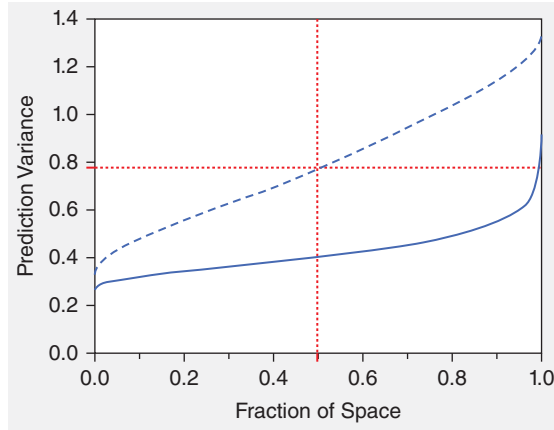
| Run | $X_1$ | $X_2$ | $X_3$ | $X_4$ |
|-----|-------|-------|-------|-------|
| 1   | 0.6   | 0.27  | 0.1   | 0.03  |
| 2   | 0.5   | 0.37  | 0.1   | 0.03  |
| 3   | 0.4   | 0.47  | 0.1   | 0.03  |
| 4   | 0.6   | 0.1   | 0.27  | 0.03  |
| 5   | 0.4   | 0.285 | 0.285 | 0.03  |
| 6   | 0.5   | 0.1   | 0.37  | 0.03  |
| 7   | 0.4   | 0.1   | 0.47  | 0.03  |
| 8   | 0.6   | 0.1   | 0.245 | 0.055 |
| 9   | 0.4   | 0.1   | 0.445 | 0.055 |
| 10  | 0.6   | 0.22  | 0.1   | 0.08  |
| 11  | 0.5   | 0.32  | 0.1   | 0.08  |
| 12  | 0.4   | 0.42  | 0.1   | 0.08  |
| 13  | 0.6   | 0.1   | 0.22  | 0.08  |
| 14  | 0.4   | 0.26  | 0.26  | 0.08  |
| 15  | 0.4   | 0.1   | 0.42  | 0.08  |

**Other Design Construction Methods for Constrained Mixture Experiments** There are two other design construction techniques that can be useful for constrained mixture experiments. The **distance-based** method attempts to spread the design points out uniformly over the feasible region. The criterion does not depend on the model that has been assumed, it strives to spread the points out uniformly. The algorithm is simple: Start with the point that is as close as possible to a vertex of the unconstrained region, and then add the point for which the Euclidean distance to the first point is a maximum. All subsequent points are added similarly; that is, choose the point for which the minimum Euclidian distance to the other points in the design is maximized. This method tends to fill the feasible region with

**TABLE 13.9 A 15-Run *I*-Optimal Design for the Flair Problem**

| Run | $X_1$ | $X_2$ | $X_3$ | $X_4$ |
|-----|-------|-------|-------|-------|
| 1   | 0.4   | 0.1   | 0.47  | 0.03  |
| 2   | 0.4   | 0.42  | 0.1   | 0.08  |
| 3   | 0.4   | 0.1   | 0.42  | 0.08  |
| 4   | 0.5   | 0.37  | 0.1   | 0.03  |
| 5   | 0.5   | 0.32  | 0.1   | 0.08  |
| 6   | 0.5   | 0.1   | 0.37  | 0.03  |
| 7   | 0.6   | 0.1   | 0.22  | 0.08  |
| 8   | 0.6   | 0.1   | 0.245 | 0.055 |
| 9   | 0.6   | 0.22  | 0.1   | 0.08  |
| 10  | 0.4   | 0.26  | 0.26  | 0.08  |
| 11  | 0.6   | 0.27  | 0.1   | 0.03  |
| 12  | 0.4   | 0.1   | 0.445 | 0.055 |
| 13  | 0.4   | 0.47  | 0.1   | 0.03  |
| 14  | 0.4   | 0.285 | 0.285 | 0.03  |
| 15  | 0.6   | 0.1   | 0.27  | 0.03  |





**Figure 13.16** Fraction of design space plot for the two experiments for the Flair problem. The dashed line is the extreme vertices design and the solid line is the  $I$ -optimal design.

points and usually forces more runs into the interior of the region compared to either the  $D$ - or  $I$ -optimal approach. Various implementations of this method can use either a set of candidate starting points or a coordinate exchange approach.

To illustrate a distance-based design, consider the following mixture problem:

$$\begin{aligned}x_1 + x_2 + x_3 &= 1 \\0.1 &\leq x_1 \leq 0.8 \\0 &\leq x_2 \leq 0.75 \\0 &\leq x_3 \leq 0.6\end{aligned}$$

A 15-run distance-based design constructed using Design-Expert is shown in Table 13.10. A point exchange approach was used to construct this design, using the vertices, centers of edges, axial blends, the overall centroid and some additional interior check blends as the candidate points.

Another approach that can be used to create designs that essentially fill the feasible region is the **space-filling** concept that is used in many computer experiments. JMP implements this approach for mixtures, and a 15-run space-filling design for the mixture problem above is shown in Table 13.11.

Figures 13.17 and 13.18 show the ternary plots for the two designs in Tables 13.10 and 13.11, respectively. Both designs do a reasonable job of filling the region. The distance-based design only has 11 distinct design points because it was constructed to have 4 replicate runs. For comparison purposes, we also constructed a 15-run  $I$ -optimal design for this problem. This design is shown in Table 13.12 and the ternary plot is shown in Figure 13.19. Figure 13.20 is a fraction of design space comparing all three designs for this problem. The distance-based design performs the worst in terms of prediction variance over the region. Interestingly, the space-filling design is slightly better than the  $I$ -optimal design for this problem. This is likely due to the number of points in the middle of the design region and their more nearly even spacing. Generally, we would expect the space-filling approach to be a good competitor to the  $I$ -optimal design in many situations. However, we

**TABLE 13.10 A 15-Run Distance-Based Design**

| Run | $X_1$    | $X_2$    | $X_3$    |
|-----|----------|----------|----------|
| 1   | 0.1      | 0.75     | 0.15     |
| 2   | 0.8      | 0.2      | 0        |
| 3   | 0.1      | 0.3      | 0.6      |
| 4   | 0.8      | 0        | 0.2      |
| 5   | 0.1      | 0.75     | 0.15     |
| 6   | 0.4      | 0        | 0.6      |
| 7   | 0.8      | 0        | 0.2      |
| 8   | 0.504167 | 0.166667 | 0.329167 |
| 9   | 0.25     | 0.15     | 0.6      |
| 10  | 0.4      | 0        | 0.6      |
| 11  | 0.254167 | 0.429167 | 0.316667 |
| 12  | 0.604167 | 0.266667 | 0.129167 |
| 13  | 0.1      | 0.3      | 0.6      |
| 14  | 0.525    | 0.475    | 0        |
| 15  | 0.329167 | 0.541667 | 0.129167 |

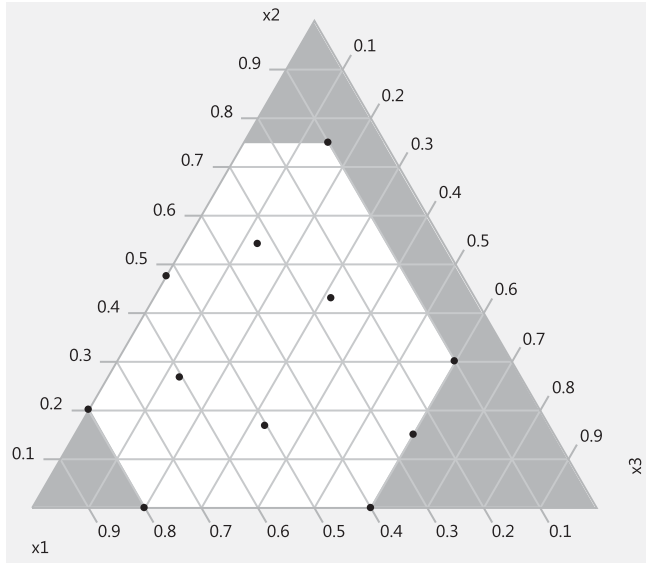
feel that in constrained mixture problems it is a good practice to construct and evaluate several designs before deciding on the final choice of design for the problem.

**Example 13.2 The Shampoo Foam Experiment** Consider the shampoo foam experiment described previously. Recall that

$$\begin{aligned}x_1 + x_2 + x_3 &= 0.5 \\0.2 &\leq x_1 \leq 0.3 \\0.07 &\leq x_2 \leq 0.10 \\0.13 &\leq x_3 \leq 0.20\end{aligned}$$

**TABLE 13.11 A 15-Run Space-Filling Design**

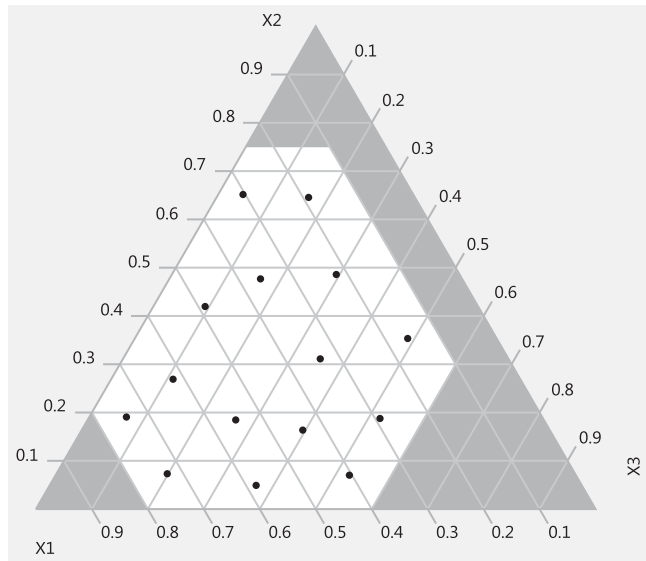
| Run | $X_1$        | $X_2$        | $X_3$        |
|-----|--------------|--------------|--------------|
| 1   | 0.4406387334 | 0.1625560436 | 0.396805223  |
| 2   | 0.7409480166 | 0.1905063539 | 0.0685456295 |
| 3   | 0.2921591673 | 0.1857187057 | 0.522122127  |
| 4   | 0.1591094659 | 0.3531195053 | 0.4877710288 |
| 5   | 0.3035483345 | 0.649635294  | 0.0468163715 |
| 6   | 0.4038409099 | 0.0681416985 | 0.5280173916 |
| 7   | 0.7276481003 | 0.0718579599 | 0.2004939398 |
| 8   | 0.2195103003 | 0.4832164184 | 0.2972732813 |
| 9   | 0.3356773509 | 0.3092710226 | 0.3550516265 |
| 10  | 0.4870841756 | 0.4174628314 | 0.095452993  |
| 11  | 0.1909472412 | 0.6432000532 | 0.1658527056 |
| 12  | 0.5796111565 | 0.0488024821 | 0.3715863614 |
| 13  | 0.5517950006 | 0.1820017602 | 0.2662032392 |
| 14  | 0.6204778129 | 0.2679798691 | 0.1115423181 |
| 15  | 0.3588797717 | 0.4749779443 | 0.1661422839 |



**Figure 13.17** Ternary plot for the distance-based design in Table 13.10.

Figure 13.14 shows the feasible region for this experiment (with the component proportions defined as  $x_i/0.5$ ). The experimenter decided to use a quadratic mixture model and a *D*-optimal design. They decided on a design with nine distinct runs and four replicates. The resulting 13-run design is shown in Table 13.13 and Figure 13.21 The figure is drawn using the *L*-pseudocomponents. The four vertices are the replicated runs.

Table 13.14 shows the details of model fitting. Based on the relatively small value of PRESS, the adjusted  $R^2$ , and the *P*-value for the special cubic term, the experimenters



**Figure 13.18** Ternary plot for the space-filling design in Table 13.10.

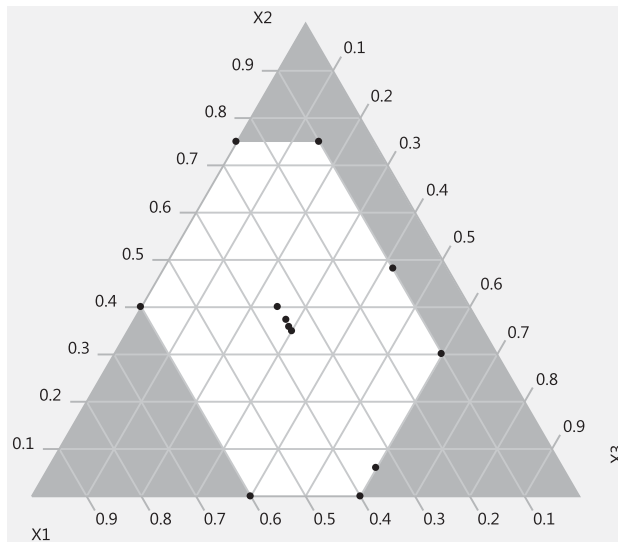
**TABLE 13.12** A 15-Run *I*-Optimal Design

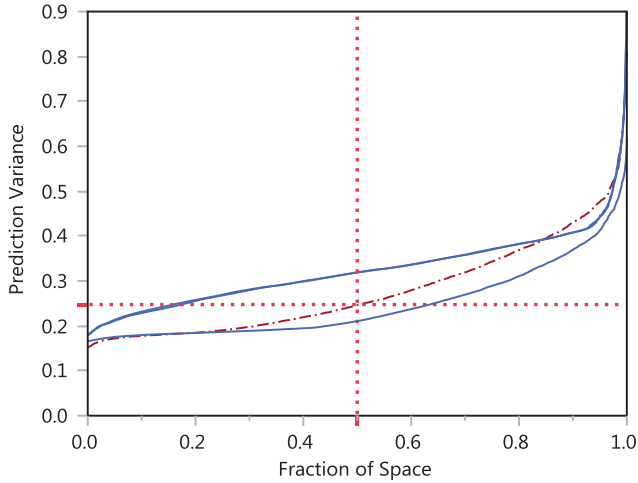
| Run | $X_1$        | $X_2$        | $X_3$        |
|-----|--------------|--------------|--------------|
| 1   | 0.6          | 0.4          | 0            |
| 2   | 0.6          | 0            | 0.4          |
| 3   | 0.35         | 0.3721027428 | 0.2778972572 |
| 4   | 0.35         | 0.35         | 0.3          |
| 5   | 0.34         | 0.06         | 0.6          |
| 6   | 0.1          | 0.75         | 0.15         |
| 7   | 0.35         | 0.4          | 0.25         |
| 8   | 0.35         | 0.4          | 0.25         |
| 9   | 0.4          | 0            | 0.6          |
| 10  | 0.1          | 0.48         | 0.42         |
| 11  | 0.1          | 0.3          | 0.6          |
| 12  | 0.25         | 0.75         | 0            |
| 13  | 0.35         | 0.3579984392 | 0.2920015608 |
| 14  | 0.6315312765 | 0.09649841   | 0.2719703135 |
| 15  | 0.0612139911 | 0.0649104104 | 0.8738755985 |

selected the special cubic model as the best model for foam height. In terms of the pseudo-components, this model is

$$\hat{y} = 146.74X_1 - 370.32X_2 + 94.90X_3 \\ + 745.43X_1X_2 + 183.21X_1X_3 + 876.64X_2X_3 - 524.99X_1X_2X_3$$

The analysis of variance for this model is shown in Table 13.15. Notice that all of the second-order terms and the special cubic term are statistically significant (their *P*-values are all less than 0.05). Residual plots from the model were satisfactory, and thus the special cubic equation above is the final model.

**Figure 13.19** Ternary plot for the *I*-optimal design in Table 13.11.



**Figure 13.20** Fraction of design spaced plots for the distance-based design in Table 13.11 (upper line), the space-filling design in Table 13.11 (lower line), and the *I*-optimal design in Table 13.12 (central line).

Figure 13.22 shows the shampoo foam height response surface, both as a contour plot and as a three-dimensional surface plot. The experimenters’ objective was to formulate a product with foam height in excess of 170 mm. From inspection of the response surface plots we note that there are many formulations that would satisfy this objective.

Figure 13.23 is the response surface trace plot. In this plot, the centroid of the constrained region is the reference mixture, and the response trace is computed along a line connecting the centroid to the original vertex  $x_i = 1$ . This is often called **Cox’s direction** (see Section 13.4 for more discussion). All three components in this trace plot exhibit strong nonlinear effects. Because all of the higher-order terms in the model were significant, this result is not unexpected.

**TABLE 13.13** *D*-Optimal Design for the Shampoo Foam Experiment, Example 13.2

| Observation | Run Order | $X_1$ | $X_2$ | $X_3$ | $y_1$ (Height, mm) |
|-------------|-----------|-------|-------|-------|--------------------|
| 1           | 11        | 1.00  | 0.000 | 0.000 | 152                |
| 2           | 12        | 1.00  | 0.000 | 0.000 | 140                |
| 3           | 3         | 0.70  | 0.300 | 0.000 | 150                |
| 4           | 6         | 0.70  | 0.300 | 0.000 | 145                |
| 5           | 5         | 0.00  | 0.300 | 0.700 | 141                |
| 6           | 2         | 0.00  | 0.300 | 0.700 | 138                |
| 7           | 10        | 0.30  | 0.000 | 0.700 | 153                |
| 8           | 4         | 0.30  | 0.000 | 0.700 | 147                |
| 9           | 8         | 0.85  | 0.150 | 0.000 | 165                |
| 10          | 7         | 0.65  | 0.000 | 0.350 | 170                |
| 11          | 1         | 0.35  | 0.300 | 0.350 | 148                |
| 12          | 13        | 0.75  | 0.075 | 0.175 | 175                |
| 13          | 9         | 0.40  | 0.075 | 0.525 | 163                |

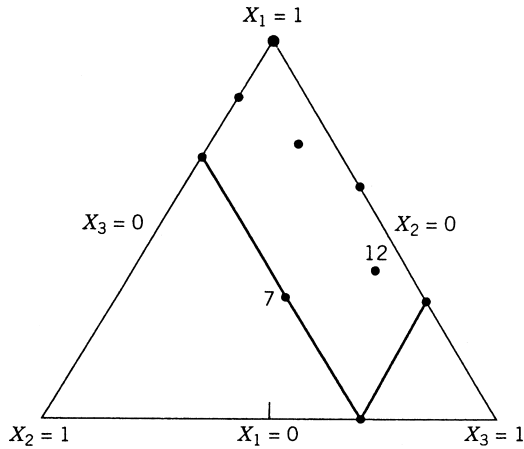


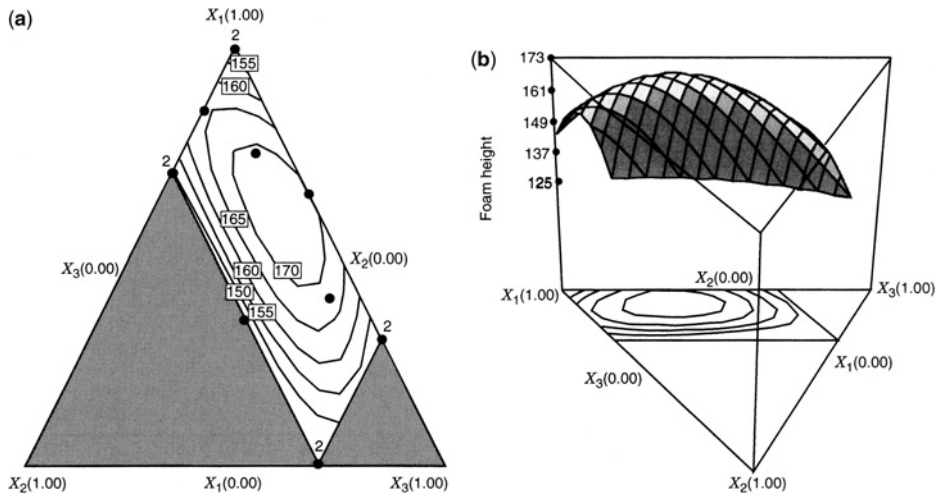
Figure 13.21 D-Optimal design for the shampoo foam experiment.

TABLE 13.14 Model Fitting for the Shampoo Foam Experiment, Example 13.2

| <i>Sequential Model Sum of Squares</i>        |                |                             |                         |           |                    |         |
|---|----------------|-----------------------------|-------------------------|-----------|--------------------|---------|
| Source  | Sum of Squares | Degrees of Freedom          | Mean Square             | F-Value   | Probability > F    |         |
| Mean  | 303,705.3      | 1                           | 303,705.3               |           |                    |         |
| Linear  | 377.4          | 2                           | 188.7                   | 1.438     | 0.2825             |         |
| Quadratic                                     | 1008.0         | 3                           | 336.0                   | 7.730     | 0.0126             |         |
| Special cubic                                 | 153.4          | 1                           | 153.4                   | 6.103     | 0.0484             |         |
| Residual                                      | 150.8          | 6                           | 25.1                    |           |                    |         |
| Lack of fit                                   | (43.8)         | 2                           | 21.9                    | 0.8193    | 0.5033             |         |
| Pure error                                    | (107.0)        | 4                           | 26.8                    |           |                    |         |
| Total   | 305,395.0      | 13                          |                         |           |                    |         |
| <i>Lack-of-Fit Tests</i>                      |                |                             |                         |           |                    |         |
| Model   | Sum of Squares | Degrees of Freedom          | Mean Square             | F-Value   | Probability > F    |         |
| Linear  | 1205.3         | 6                           | 200.9                   | 7.509     | 0.0357             |         |
| Quadratic                                     | 197.3          | 3                           | 65.8                    | 2.458     | 0.2026             |         |
| Special cubic                                 | 43.8           | 2                           | 21.9                    | 0.8193    | 0.5033             |         |
| Pure error                                    | 107.0          | 4                           | 26.8                    |           |                    |         |
| <i>ANOVA Summary Statistics of Models Fit</i> |                |                             |                         |           |                    |         |
| Source  | Unalised Terms | Residual Degrees of Freedom | Root Mean Squared Error | R-Squared | Adjusted R-Squared | PRESS   |
| Linear  | 3              | 10                          | 11.46                   | 0.2234    | 0.0681             | 2054.97 |
| Quadratic                                     | 6              | 7                           | 6.59                    | 0.8199    | 0.6913             | 1035.39 |
| Special cubic                                 | 7              | 6                           | 5.01                    | 0.9107    | 0.8215             | 657.08  |

**TABLE 13.15 ANOVA for the Special Cubic Model, Shampoo Foam Experiment**

| Source                   | Sum of Squares       | Degrees of Freedom                | Mean Square    | F-Value                              | Probability > F          |
|--------------------------|----------------------|-----------------------------------|----------------|--------------------------------------|--------------------------|
| Model                    | 1538.9               | 6                                 | 256.48         | 10.20                                | 0.0062                   |
| Residual                 | 150.8                | 6                                 | 25.14          |                                      |                          |
| Lack of fit              | (43.8)               | 2                                 | 21.92          | 0.8193                               | 0.5033                   |
| Pure error               | (107.0)              | 4                                 | 26.75          |                                      |                          |
| Corrected total          | 1689.7               | 12                                |                |                                      |                          |
| Root means squared error | 5.01                 | <i>R</i> -Squared 0.9107          |                |                                      |                          |
| Dependent mean           | 152.85               | Adjusted <i>R</i> -squared 0.8215 |                |                                      |                          |
| Coefficient of variation | 3.28%                |                                   |                |                                      |                          |
| Independent Variable     | Coefficient Estimate | Degrees of Freedom                | Standard Error | <i>t</i> for $H_0$ : Coefficient = 0 | Probability >   <i>t</i> |
| $X_1$                    | 146.74               | 1                                 | 3.49           | Not applicable                       |                          |
| $X_2$                    | -370.32              | 1                                 | 130.14         | Not applicable                       |                          |
| $X_3$                    | 94.90                | 1                                 | 15.01          | Not applicable                       |                          |
| $X_1X_2$                 | 745.43               | 1                                 | 185.13         | 4.026                                | 0.0069                   |
| $X_1X_3$                 | 183.21               | 1                                 | 43.18          | 4.243                                | 0.0054                   |
| $X_2X_3$                 | 876.64               | 1                                 | 187.33         | 4.680                                | 0.0034                   |
| $X_1X_2X_3$              | -524.99              | 1                                 | 212.51         | -2.470                               | 0.0484                   |



**Figure 13.22** The shampoo foam height response surface, (a) Contour plot. (b) Three-dimensional surface plot. The solid dots in the contour plot are the design points.

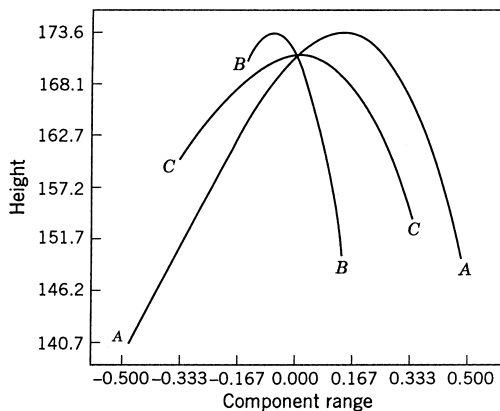


Figure 13.23 Response surface trace plot for the shampoo foam experiment.

### 13.1.4 Multicomponent Constraints

Sometimes in addition to upper and/or lower bounds on the individual component proportions, there are other linear constraints, such as

$$C_j \leq A_{1j}x_1 + A_{2j}x_2 + \cdots + A_{qj}x_q \leq D_j \quad (13.9)$$

for  $j=1, 2, \dots, m$ . These types of constraints are called **multicomponent constraints**. Some of the constants  $A_{ij}$  may be zero, so that in general, not all components appear in each constraint.

When multicomponent constraints are present, they can change the shape of the feasible region for the experiment. That is, they can change the number of vertices, edges, and so on, that define the feasible space. To illustrate, consider the following example:

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ 0.1 &\leq x_1 \\ x_2 &\leq 0.8 \\ x_3 &\leq 0.7 \\ 0.4 &\leq x_1 + 0.8x_2 \end{aligned}$$

Figure 13.24 shows the feasible region for this design. Notice that the effect of the multicomponent constraint is to cut off part of the region that is defined by the upper- and lower-bound constraints. As a result, both the number of vertices and the number of edges in the feasible region are increased by one.

Piepel (1988) published an algorithm called CONVRT, which can locate the extreme vertices of the region when multi-component constraints are active. Both Design-Expert and JMP can handle multicomponent constraints, and generate appropriate designs for these oddly-shaped regions.



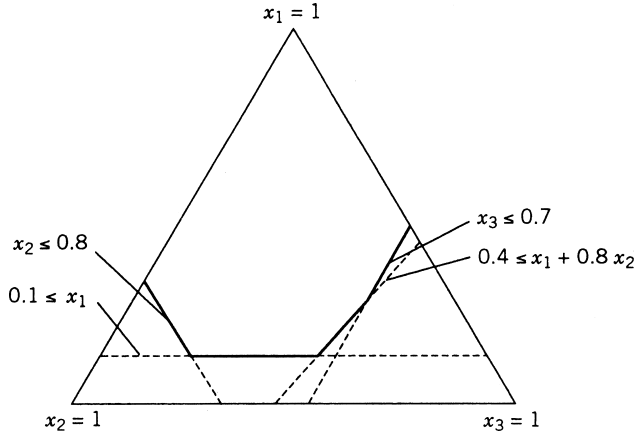


Figure 13.24 Three-component constrained region.

13.2 MIXTURE EXPERIMENTS USING RATIOS OF COMPONENTS

In some mixture experiments, the experimenter may wish to work with **ratios** of the mixture components instead of the original component proportions. For example, in formulating a detergent, it may be convenient or natural for the chemist to think of the ratio of an emulsifier to water rather than the actual proportion of each ingredient in a blend. In such an experiment, it might be natural to think of each ingredient as the ratio of the ingredient to water.

There are usually several ways to set up the ratios. For example, if there are three components, then one way to define the ratio variables is

$$r_1 = \frac{x_1}{x_3}, \quad r_2 = \frac{x_2}{x_3}$$

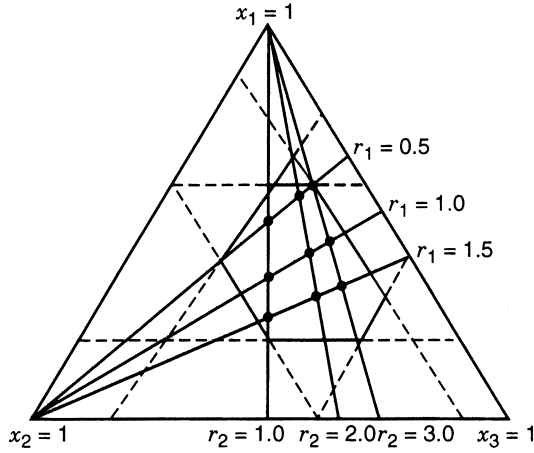
Another possibility is

$$r_1 = \frac{x_2}{x_1}, \quad r_2 = \frac{x_2}{x_3}$$

In general, each ratio  $r_1$  or  $r_2$  must contain at least one of the components used in the other ratio. This holds true in general: If there are  $q$  components, then we may define  $q-1$  ratios  $r_1, r_2, \dots, r_{q-1}$ , and each ratio should contain at least one of the components used in at least one of the other ratios in the set. The  $q-1$  ratio variables are *independent*; consequently any type of standard response surface polynomial can be fitted to the ratios  $r_1, r_2, \dots, r_{q-1}$ . Obviously, standard response surface designs can be used in these situations.

**Example 13.3 Gasoline Blending** Cornell (2002) presents a three-component example of blending refinery products to form gasoline. The response variable of interest is octane number. The ratios used are

$$r_1 = \frac{x_3}{x_1}, \quad r_2 = \frac{x_3}{x_2}$$



**Figure 13.25** Feasible design region for Example 12.3, along with design points in the ratio variables.

and the process currently operates at  $r_1 = 1, r_2 = 2$  and at a composition of  $x_1 = 0.4, x_2 = 0.2,$  and  $x_3 = 0.4$ . The region of interest is constrained by

$$\begin{aligned} 0.2 &\leq x_1 \leq 0.6 \\ 0.1 &\leq x_2 \leq 0.4 \\ 0.2 &\leq x_3 \leq 0.6 \end{aligned}$$

The feasible region is shown in Fig. 13.25. Note that for these models, the constraints are still specified in terms of the proportions of each ingredient. Hence the experimenter should track both proportions and ratios as the design is being constructed.

Suppose that the experimenter wishes to fit the quadratic model

$$y = \beta_0 + \beta_1 r_1 + \beta_2 r_2 + \beta_{12} r_1 r_2 + \beta_{11} r_1^2 + \beta_{22} r_2^2 + \varepsilon$$

in the ratio variables. Any type of response surface design in the two variables  $r_1$  and  $r_2$  could be used. Cornell chose values of the ratios equal to  $r_1 = (0.5, 1.0, 1.5)$  and  $r_2 = (1.0, 2.0, 3.0)$  and used a  $3^2$  design. The rays representing these ratios are shown in Fig. 13.25. The nine design points are at the intersections of the six rays. Table 13.16 shows the design in both the ratios and the original component proportions. The settings of the original components in terms of the ratios (which are required to run the experiment) were found by solving

$$r_1 = \frac{x_3}{x_1}, \quad r_2 = \frac{x_3}{x_2}, \quad x_1 + x_2 + x_3 = 1$$

**TABLE 13.16 Octane Data for Example 13.3**

| Run | $r_1 = x_3/x_1$ | $r_2 = x_3/x_2$ | $x_1$ | $x_2$ | $x_3$ | Octane Number $y$ |
|-----|-----------------|-----------------|-------|-------|-------|-------------------|
| 1   | 0.5             | 1.0             | 0.50  | 0.25  | 0.25  | 82.8              |
| 2   | 1.0             | 1.0             | 0.33  | 0.33  | 0.33  | 87.3              |
| 3   | 1.5             | 1.0             | 0.25  | 0.37  | 0.37  | 84.9              |
| 4   | 0.5             | 2.0             | 0.57  | 0.14  | 0.29  | 85.6              |
| 5   | 1.0             | 2.0             | 0.40  | 0.20  | 0.40  | 93.0              |
| 6   | 1.5             | 2.0             | 0.31  | 0.23  | 0.46  | 89.3              |
| 7   | 0.5             | 3.0             | 0.60  | 0.10  | 0.30  | 87.2              |
| 8   | 1.0             | 3.0             | 0.43  | 0.14  | 0.43  | 89.6              |
| 9   | 1.5             | 3.0             | 0.33  | 0.17  | 0.50  | 90.6              |

for  $x_1, x_2,$  and  $x_3,$  yielding

$$x_1 = \frac{r_2}{r_1 + r_1 r_2 + r_2}$$

$$x_2 = \frac{r_1}{r_1 + r_1 r_2 + r_2}$$

$$x_3 = \frac{r_1 r_2}{r_1 + r_1 r_2 + r_2}$$

Table 13.17 presents the analysis of variance summary for the quadratic model. The model coefficients in this table are given in terms of **coded** orthogonal variables  $z_1$  and  $z_2,$  where

$$z_1 = \frac{r_1 - 1.0}{0.5}, \quad z_2 = \frac{r_2 - 2.0}{1.0}$$

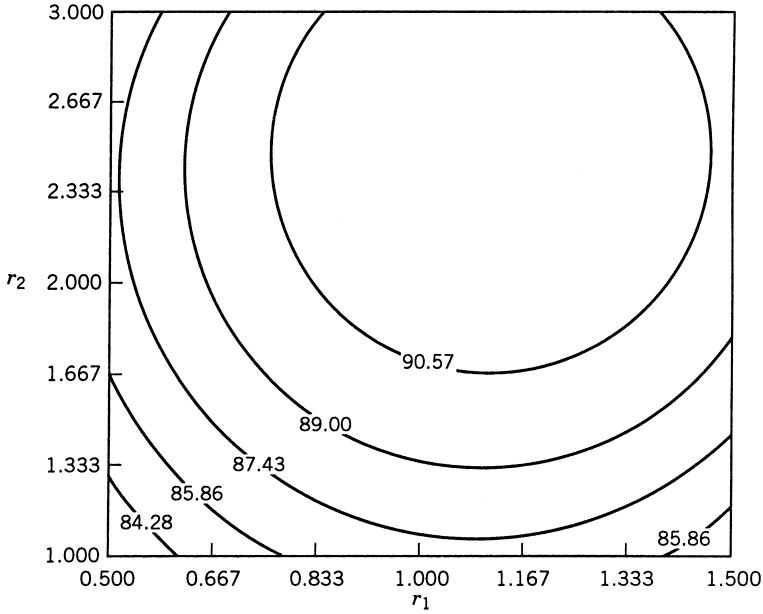
**TABLE 13.17 Analysis of Variance for the Quadratic Model in Example 12.3**

| Source          | Sum of Squares | Degrees of Freedom | Mean Square | $F$ -Value | Probability $> F$ |
|-----------------|----------------|--------------------|-------------|------------|-------------------|
| Model           | 71.04          | 5                  | 14.208      | 5.205      | 0.1025            |
| Residual        | 8.19           | 3                  | 2.730       |            |                   |
| Corrected total | 79.23          | 8                  |             |            |                   |

Root mean square error 1.652  
 Dependent mean 87.811  
 Coefficient of variation 1.88%

$R$ -squared 0.8966  
 Adjusted  $R$ -squared 0.7244

| Independent Variable | Coefficient Estimate | Degrees of Freedom | Standard Error | $t$ for $H_0$ Coefficient = 0 | Probability $>  t $ |
|----------------------|----------------------|--------------------|----------------|-------------------------------|---------------------|
| Intercept            | 91.456               | 1                  | 1.231          | 74.27                         |                     |
| $A = z_1$            | 1.533                | 1                  | 0.674          | 2.273                         | 0.1076              |
| $B = z_2$            | 2.067                | 1                  | 0.674          | 3.064                         | 0.0548              |
| $A^2$                | -3.233               | 1                  | 1.168          | -2.768                        | 0.0697              |
| $B^2$                | -2.233               | 1                  | 1.168          | -1.912                        | 0.1519              |
| $AB$                 | 0.325                | 1                  | 0.826          | 0.3934                        | 0.7203              |



**Figure 13.26** Octane contour plot in terms of the ratios  $r_1$  and  $r_2$ , Example 12.3.

In terms of these variables, the model is

$$\hat{y} = 91.456 + 1.533z_1 + 2.067z_2 + 0.325z_1z_2 - 3.233z_1^2 - 2.233z_2^2$$

The corresponding model in terms of the ratios  $r_1$  and  $r_2$  is

$$\hat{y} = 63.689 + 27.633r_1 + 10.350r_2 + 0.650r_1r_2 - 12.933r_1^2 - 2.233r_2^2$$

Figure 13.26 is the octane contour plot in terms of the ratios  $r_1$  and  $r_2$ . If the experimenter wants to blend gasoline with an octane number of 89, we see from this plot that several blends will satisfy this objective. For example,  $r_1 = 1.0$  and  $r_2 = 1.33$  is a feasible set of ratios. In terms of the original components, this point is

$$x_1 = \frac{r_2}{r_1 + r_1r_2 + r_2} = \frac{1.33}{1.0 + (1.0)1.33 + 1.33} = 0.36$$

$$x_2 = \frac{r_1r_2}{r_1 + r_1r_2 + r_2} + \frac{(1.0)1.33}{1.0 + (1.0)1.33 + 1.33} = 0.36$$

$$x_3 = \frac{r_1}{r_1 + r_1r_2 + r_2} = \frac{1.0}{1.0 + (1.0)1.33 + 1.33} = 0.28$$

The ratio variables approach is a very useful one, and we recommend it in cases where it is convenient or natural for the experimenters to work in terms of these variables. The approach does have two potential disadvantages. First, the interpretation is in terms of a function of the original component proportions (the ratios), and this may not be a natural

framework for the experimenter to work in. The second disadvantage relates to how the design in terms of the ratio variables fits into the feasible mixture space. By referring to Fig. 13.25, we see that the feasible mixture space for Example 13.3 is actually larger than the region of experimentation for the  $3^2$  design in the ratios  $r_1$  and  $r_2$ . If it is important to predict the response over this entire region, then this  $3^2$  design would not be as a good a choice as a  $D$ -optimal or an  $I$ -optimal design in the original component proportions. See Piepel and Cornell (1994) for other comments on the use of ratios.

### 13.3 PROCESS VARIABLES IN MIXTURE EXPERIMENTS

#### 13.3.1 Mixture-Process Model and Design Basics

**Process variables** are factors in an experiment that are not mixture components but could affect the blending properties of the mixture ingredients. For example, the effectiveness of an etching solution (measured as an etch rate, say) is not only a function of the proportions of the three acids that are combined to form the mixture, but also depends on the temperature of the solution and the agitation rate.

In general, there may be  $q$  mixture components  $x_1, x_2, \dots, x_q$  and  $p$  process variables  $z_1, z_2, \dots, z_p$ . The etching example above has  $q = 3$  mixture components and  $p = 2$  process variables. The reason we consider this as a separate class of experiment is that the mixture components are interdependent since they are constrained to sum to one, while the process variables can be varied independently.

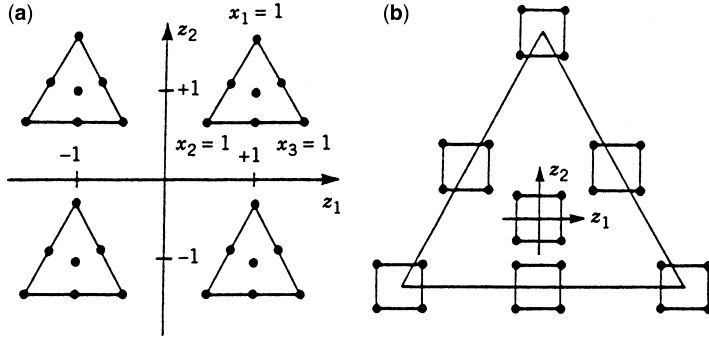
There are two approaches to including process variables in mixture experiments. The first approach involves transforming the mixture variables into  $q-1$  independent variables and then treating all  $p+q-1$  factors in a conventional (nonmixture) experiment. Using ratios of the components is one transformation of the mixture variables that is widely used in practice. To illustrate, consider the etching solution described above, and let the ratios be defined as

$$r_1 = \frac{x_1}{x_3}, \quad r_2 = \frac{x_2}{x_3}$$

The process variables are  $z_1 = \text{temperature}$  and  $z_2 = \text{agitation rate}$ . Then a simple model relating these factors is

$$\begin{aligned} E(y) = & \beta_0 + \beta_1 r_1 + \beta_2 r_2 + \alpha_1 z_1 + \alpha_2 z_2 \\ & + \delta_{11} r_1 z_1 + \delta_{12} r_1 z_2 + \delta_{21} r_2 z_1 + \delta_{22} r_2 z_2 \\ & + \beta_{12} r_1 r_2 + \alpha_{12} z_1 z_2 \end{aligned}$$

In this model, we may obtain information about the linear and nonlinear blending properties of the mixture components through the parameters  $\beta_1$ ,  $\beta_2$ , and  $\beta_{12}$ ; information about the main effects and interaction of the process variables through the parameters  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_{12}$ ; and information on the relation between the mixture components and the process variables through the parameters  $\delta_{ij}$ . By using the ratios the proportions of the ingredients have been expressed without constraints, and hence each of  $r_1$ ,  $r_2$ ,  $z_1$ , and  $z_2$  can be varied independently when choosing a design. In addition to the ratio approach, there are other



**Figure 13.27** Two different presentations of a 28-point combined design for three mixture components and two process variables. (a) The seven-point simplex centroid constructed at each of the  $2^2 = 4$  points of the factorial arrangement. (b) A complete  $2^2$  constructed at each of the seven points of the simplex centroid design. [From Cornell (2002), with permission of the publisher.]

transformations that could be used to convert the mixture components into a set of  $q-1$  independent mixture-related variables. For a good discussion of these transformations, see Cornell (2002).

The second approach is to model the mixture components directly. This involves setting up a mixture design at each treatment combination of the factorial experiment used for the process variables. Alternatively, we can think of this as setting up a factorial in the process variables at each point in the mixture design. These two viewpoints are shown in Fig. 13.27 for the etching problem described above, assuming that a  $2^2$  factorial is used for the process variables and a simplex-centroid design is used for the three mixture components. The combined design has 28 points, and can be used to collect response data for fitting a **combined model** that includes both the mixture and the process variables. If we wish to fit a quadratic model in the mixture variables and a model with both main effects and the two-factor interaction in the process variables, then the combined model can be written as

$$\begin{aligned}
 E(y) = & \sum_{i=1}^3 \beta_i x_i + \sum_{i < j=2}^2 \sum_{j=2}^3 \beta_{ij} x_i x_j \\
 & + \sum_{k=1}^2 \left[ \sum_{i=1}^3 \alpha_{ik} x_i + \sum_{i < j=2}^3 \sum_{j=2}^3 \alpha_{ijk} x_i x_j \right] z_k \\
 & + \left[ \sum_{i=1}^3 \delta_{i12} x_i + \sum_{i < j=2}^3 \sum_{j=2}^3 \delta_{ij12} x_i x_j \right] z_1 z_2 \tag{13.10}
 \end{aligned}$$

This model has 24 parameters. Fortunately, they have relatively straightforward interpretations. The first six terms on the right-hand side of Equation 13.10 are the linear and nonlinear blending portions of the model. They do not depend on the settings of the process variables. All remaining parameters measure the effects of the process variables on the linear and nonlinear blending properties of the mixture components. Note that terms involving just the process variables alone can not be included in the model, since the constraint on

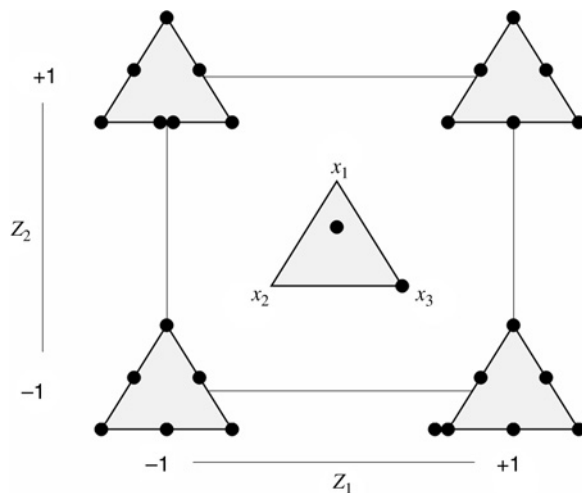
the mixture components restrict their inclusion. See Anderson-Cook et al. (2004) for more details.

The usual approach to analyzing such a model is to fit the complete combined model and then possibly reduce the model by testing hypotheses on individual parameters or groups of parameters, using the extra-sum-of-squares method. This model would be very useful for predicting the response of a particular mixture formulation at combinations of the process variables that are not factorial design points. To predict at design points in the factorial space one could also fit individual mixture models to the mixture experiments at each factorial point. Contour plotting of the response in terms of the mixture components at each point in the factorial array is then a reasonably effective way to interpret the response surface.

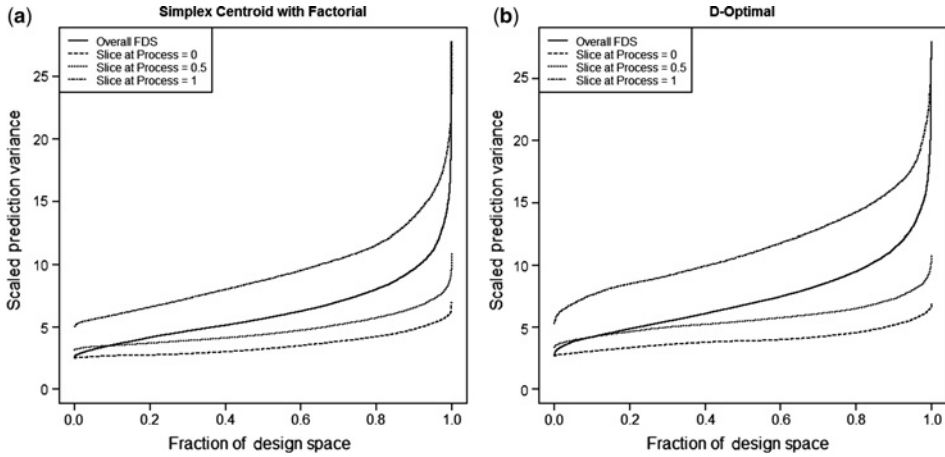
As the numbers of mixture components and process variables increase, the size of these combined designs becomes very large. Several authors have suggested fractionating these combined designs, although there are no obvious best choices of the fractional design. This is a significant problem when there are constraints on the mixture components and the mixture space is a hyperpolytope. Either the *D*- or the *I*-optimality criteria could be used for generating a design for these problems. Goos and Donev (2007) provide additional details about design construction for mixture process experiments. In order to produce valid models, we caution the experimenter to be sure that there are sufficient runs in the mixture part of the experiment to support the mixture polynomial at each treatment combination in the factorial design. That is, if the combined design has too many runs, we would prefer to fractionate the process variable factorial portion of the design.

Figure 13.28 shows a design created using the *D*-optimality criterion in Design-Expert where the experimenter specified the model of Equation 13.10, and a total of 28 runs, where 2 runs were allocated for estimating lack of fit and 2 for pure error. The addition of the runs for  $(z_1, z_2)=(0, 0)$  are those designated for assessing lack of fit.

As with other designs, variance dispersion graphs (VDGs) and fraction of design space (FDS) plots can be used to compare the prediction performance of mixture-process variables. Because we expect the performance to change differently as we move in the mixture space compared to moving in the process space, Goldfarb et al. (2004a, 2004b) developed



**Figure 13.28** *D*-optimal design generated by Design-Expert for model in Equation 13.10.



**Figure 13.29** Fraction of design space plot for designs shown in Figs. 13.27 and 13.30. (a) Simplex centroid with factorial. (b) *D*-optimal.

methods for examining these features separately. Figure 13.29 shows an FDS plot for the designs considered in Figs. 13.27 and 13.28. The additional lines for each plot show an FDS curve for various slices of the process space.

From the FDS plots, we can see that both designs have similar maxima and minima for SPV. Since the curves for different fixed process shrinkages are relatively flat and the different slices are separated from each other, the location in terms of the mixture components does not affect the SPV as much as the location in terms of the process variables. The *D*-optimal design has slightly flatter FDS curves for the majority of the design space, which makes it slightly more desirable for prediction performance. The addition of the replicates also allows estimation of pure error for this design.

**Example 13.4 Adhesive Formulation** An adhesive is being formulated for use in an aerospace application. The adhesive consists of a resin  $x_1$  and two crosslinkers,  $x_2$  and  $x_3$ . The mixture constraints for these variables are

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ 0.7 &\leq x_1 \leq 0.9 \\ 0.05 &\leq x_2 \leq 0.10 \\ 0.05 &\leq x_3 \leq 0.20 \end{aligned}$$

The adhesive is applied to the components, and then the entire assembly is cured for 12 hr at controlled temperature and humidity. The temperature  $z_1$  and relative humidity  $z_2$  are process variables that can be controlled by the experimenter. The ranges of these process variables that the experimenters think are appropriate are  $40^\circ\text{F} \leq \text{temperature} \leq 100^\circ\text{F}$  and  $15\% \leq \text{relative humidity} \leq 85\%$ . The response variable of interest is the pulloff force required to separate the components after curing. It should exceed 40 pounds.

This is a three-component mixture experiment with two process variables. A reasonable model for the response is the combined model in Equation 13.10, and a design strategy similar to the one illustrated in Fig. 13.27 using a  $2^2$  factorial for the process variables, is



appropriate. However, since the mixture design space is not a simplex, the experimenters cannot use a simplex-type design for the mixture components. Instead, they chose to use a *D*-optimal design, and assumed that a quadratic model would be satisfactory to represent the mixture components. The candidate set of points was taken to be the four vertices, the four edge centers, the overall centroid, and the axial check blends for the constrained mixture region at each of the four corners of the square in the process variable space. This produces a total of 52 candidate design points. The model has 24 terms, and the experimenters chose to use a design with 34 runs. There are 29 distinct design points, and five runs are replicated to provide an estimate of error. The *D*-optimal design produced by Design-Expert is shown in Table 13.18. This table contains both the actual mixture components and process variable levels, as well as their coded levels.

Table 13.19 contains the Design-Expert output that results from fitting the complete 24-term model in Equation 13.10 to the pulloff force response data. Note that there are several nonsignificant model terms. A stepwise regression algorithm was then employed, resulting in the reduced model shown in Table 13.20. There is no significant lack of fit in this model, and the residual plots are satisfactory, so we conclude that the reduced model is adequate.

Figure 13.30 presents contour plots of the predicted pulloff force at each of the four combinations of the process variables, temperature and relative humidity. The design points are shown as circles on the plot. There are many recipes for the adhesive formulation that will produce pulloff force values in excess of 40 lb, except at low temperature and high relative humidity. The highest values of the pulloff force are obtained when temperature is high and relative humidity is low. The experimenters examined contour plots at several combinations of the process variables other than the cube corners. One of these, at temperature 70°F and relative humidity 30%, is shown in Fig. 13.31. The experimenters selected these levels of the process variables for the final curing conditions because they were easy to maintain in the curing chamber. There are several combinations of the mixture components that will produce a pulloff force that meets or exceeds the requirements at these curing conditions.

### 13.3.2 Split-Plot Designs for Mixture-Process Experiments

Since the nature of the mixture and process variables are often very different, it is common for there to be randomization restrictions on one set of these inputs. By considering a split-plot structure, we can design an experiment which allows us to more efficiently explore the relationship between mixture and process factors on our response. This approach combines the mixture-process methodology with the split-plot concepts discussed in Chapters 3, 5, and 10.

For example, it is sometimes convenient to make larger batches of mixtures. Then we can subdivide these batches to expose them to different process variable conditions. In this case the mixture variables are the whole plot factors, and the process variables are the subplot factors. Alternatively, in some experimental settings making up the mixtures is not time or resource intensive. Here we may wish to mix several different sets of mixtures, and then expose them in groups to different process variable conditions. Since we are running multiple sets of mixtures for each setting of the process variables, the whole plot factors in this case are the process factors, and the mixture factors and the subplots.

TABLE 13.18 Design for Example 13.4 with Three Mixture Components and Two Process Variables

| Std | Run | Actual Mixture and Process Variables |        |        |              |             |             |             | Mixture Pseudocomponent |             |             | Process Variables |  |  |
|-----|-----|--------------------------------------|--------|--------|--------------|-------------|-------------|-------------|-------------------------|-------------|-------------|-------------------|--|--|
|     |     | A: Resin                             | B: CX1 | C: CX2 | D: Temp (°F) | E: R.H. (%) | A ( $x_1$ ) | B ( $x_2$ ) | C ( $x_3$ )             | D ( $z_1$ ) | E ( $z_2$ ) | Force.y (lb)      |  |  |
| 1   | 26  | 0.90                                 | 0.05   | 0.05   | 40           | 15          | 1.000       | 0.000       | 0.000                   | -1          | -1          | 44                |  |  |
| 2   | 6   | 0.90                                 | 0.05   | 0.05   | 100          | 15          | 1.000       | 0.000       | 0.000                   | 1           | -1          | 70                |  |  |
| 3   | 21  | 0.90                                 | 0.05   | 0.05   | 40           | 85          | 1.000       | 0.000       | 0.000                   | -1          | 1           | 19                |  |  |
| 4   | 8   | 0.90                                 | 0.05   | 0.05   | 100          | 85          | 1.000       | 0.000       | 0.000                   | 1           | 1           | 33                |  |  |
| 5   | 5   | 0.70                                 | 0.10   | 0.20   | 40           | 15          | 0.000       | 0.250       | 0.750                   | -1          | -1          | 48                |  |  |
| 6   | 2   | 0.70                                 | 0.10   | 0.20   | 100          | 15          | 0.000       | 0.250       | 0.750                   | 1           | -1          | 72                |  |  |
| 7   | 30  | 0.70                                 | 0.10   | 0.20   | 40           | 85          | 0.000       | 0.250       | 0.750                   | -1          | 1           | 32                |  |  |
| 8   | 24  | 0.70                                 | 0.10   | 0.20   | 100          | 85          | 0.000       | 0.250       | 0.750                   | 1           | 1           | 59                |  |  |
| 9   | 17  | 0.75                                 | 0.05   | 0.20   | 40           | 15          | 0.250       | 0.000       | 0.750                   | -1          | -1          | 32                |  |  |
| 10  | 23  | 0.75                                 | 0.05   | 0.20   | 100          | 15          | 0.250       | 0.000       | 0.750                   | 1           | -1          | 58                |  |  |
| 11  | 33  | 0.75                                 | 0.05   | 0.20   | 40           | 85          | 0.250       | 0.000       | 0.750                   | -1          | 1           | 21                |  |  |
| 12  | 9   | 0.75                                 | 0.05   | 0.20   | 100          | 85          | 0.250       | 0.000       | 0.750                   | 1           | 1           | 38                |  |  |
| 13  | 11  | 0.85                                 | 0.10   | 0.05   | 40           | 15          | 0.750       | 0.250       | 0.000                   | -1          | -1          | 51                |  |  |
| 14  | 20  | 0.85                                 | 0.10   | 0.05   | 100          | 15          | 0.750       | 0.250       | 0.000                   | 1           | -1          | 76                |  |  |
| 15  | 19  | 0.85                                 | 0.10   | 0.05   | 40           | 85          | 0.750       | 0.250       | 0.000                   | -1          | 1           | 22                |  |  |
| 16  | 22  | 0.85                                 | 0.10   | 0.05   | 100          | 85          | 0.750       | 0.250       | 0.000                   | 1           | 1           | 49                |  |  |
| 17  | 10  | 0.80                                 | 0.07   | 0.13   | 40           | 85          | 0.500       | 0.125       | 0.375                   | -1          | 1           | 17                |  |  |

|    |    |      |      |      |     |    |       |       |       |    |    |    |
|----|----|------|------|------|-----|----|-------|-------|-------|----|----|----|
| 18 | 16 | 0.72 | 0.07 | 0.20 | 100 | 15 | 0.125 | 0.125 | 0.750 | 1  | -1 | 69 |
| 19 | 32 | 0.72 | 0.07 | 0.20 | 40  | 15 | 0.125 | 0.125 | 0.750 | -1 | -1 | 40 |
| 20 | 3  | 0.78 | 0.10 | 0.13 | 40  | 15 | 0.375 | 0.250 | 0.375 | -1 | -1 | 37 |
| 21 | 7  | 0.80 | 0.07 | 0.13 | 100 | 85 | 0.500 | 0.125 | 0.375 | 1  | 1  | 46 |
| 22 | 29 | 0.78 | 0.10 | 0.13 | 40  | 85 | 0.375 | 0.250 | 0.375 | -1 | 1  | 21 |
| 23 | 27 | 0.78 | 0.10 | 0.13 | 100 | 15 | 0.375 | 0.250 | 0.375 | 1  | -1 | 82 |
| 24 | 34 | 0.78 | 0.10 | 0.13 | 100 | 85 | 0.375 | 0.250 | 0.375 | 1  | 1  | 43 |
| 25 | 15 | 0.82 | 0.05 | 0.13 | 40  | 15 | 0.625 | 0.000 | 0.375 | -1 | -1 | 32 |
| 26 | 1  | 0.82 | 0.05 | 0.13 | 100 | 15 | 0.625 | 0.000 | 0.375 | 1  | -1 | 60 |
| 27 | 28 | 0.85 | 0.06 | 0.09 | 40  | 85 | 0.750 | 0.063 | 0.187 | -1 | 1  | 14 |
| 28 | 4  | 0.85 | 0.06 | 0.09 | 100 | 85 | 0.750 | 0.063 | 0.187 | 1  | 1  | 38 |
| 29 | 14 | 0.83 | 0.09 | 0.09 | 40  | 15 | 0.625 | 0.187 | 0.187 | -1 | -1 | 45 |
| 30 | 13 | 0.78 | 0.10 | 0.13 | 40  | 85 | 0.375 | 0.250 | 0.375 | -1 | 1  | 18 |
| 31 | 18 | 0.72 | 0.07 | 0.20 | 100 | 15 | 0.125 | 0.125 | 0.750 | 1  | -1 | 70 |
| 32 | 25 | 0.85 | 0.10 | 0.05 | 40  | 85 | 0.750 | 0.250 | 0.000 | -1 | 1  | 10 |
| 33 | 12 | 0.78 | 0.10 | 0.13 | 100 | 85 | 0.375 | 0.250 | 0.375 | 1  | 1  | 52 |
| 34 | 31 | 0.85 | 0.10 | 0.05 | 100 | 85 | 0.750 | 0.250 | 0.000 | 1  | 1  | 42 |

**TABLE 13.19 Design-Expert Output for the Full Model, Example 13.4**

| Source                | Sum of Squares | DF       | Mean Square   | F-Value        | Prob > F      |
|-----------------------|----------------|----------|---------------|----------------|---------------|
| Model                 | 12133.68       | 23       | 527.55        | 24.40          | <0.0001       |
| <i>Linear Mixture</i> | <i>640.41</i>  | <i>2</i> | <i>320.20</i> | <i>14.81</i>   | <i>0.0010</i> |
| AB                    | 29.18          | 1        | 29.18         | 1.35           | 0.2723        |
| AC                    | 78.10          | 1        | 78.10         | 3.61           | 0.0865        |
| AD                    | 371.48         | 1        | 371.48        | 17.18          | 0.0020        |
| AE                    | 1044.33        | 1        | 1044.33       | 48.30          | <0.0001       |
| BC                    | 37.54          | 1        | 37.54         | 1.74           | 0.2170        |
| BD                    | 11.04          | 1        | 11.04         | 0.51           | 0.4912        |
| BE                    | 1.214E-003     | 1        | 1.214E-003    | 5.614E-005     | 0.9942        |
| CD                    | 31.91          | 1        | 31.91         | 1.48           | 0.2523        |
| CE                    | 8.92           | 1        | 8.92          | 0.41           | 0.5351        |
| ADB                   | 19.35          | 1        | 19.35         | 0.89           | 0.3665        |
| ABE                   | 0.55           | 1        | 0.55          | 0.026          | 0.8761        |
| ACD                   | 32.17          | 1        | 32.17         | 1.49           | 0.2505        |
| ACE                   | 3.04           | 1        | 3.04          | 0.14           | 0.7157        |
| ADE                   | 20.89          | 1        | 20.89         | 0.97           | 0.3488        |
| BCD                   | 22.32          | 1        | 22.32         | 1.03           | 0.3335        |
| BCE                   | 0.49           | 1        | 0.49          | 0.023          | 0.8836        |
| BDE                   | 4.37           | 1        | 4.37          | 0.20           | 0.6625        |
| CDE                   | 8.57           | 1        | 8.57          | 0.40           | 0.5431        |
| ABDE                  | 5.76           | 1        | 5.76          | 0.27           | 0.6171        |
| ACDE                  | 35.27          | 1        | 35.27         | 1.63           | 0.2304        |
| BCDE                  | 2.89           | 1        | 2.89          | 0.13           | 0.7221        |
| Residual              | 216.20         | 10       | 21.62         |                |               |
| <i>Lack of Fit</i>    | <i>74.20</i>   | <i>5</i> | <i>14.84</i>  | <i>0.52</i>    | <i>0.7533</i> |
| <i>Pure Error</i>     | <i>142.00</i>  | <i>5</i> | <i>28.40</i>  |                |               |
| Cor Total             | 12349.88       | 33       |               |                |               |
| Std. Dev.             | 4.65           |          |               | R-Squared      | 0.9825        |
| Mean                  | 42.94          |          |               | Adj R-Squared  | 0.9422        |
| C.V.                  | 10.83          |          |               | Pred R-Squared | N/A           |
| PRESS                 | N/A            |          |               | Adeq Precision | 16.103        |

Case(s) with leverage of 1.0000: Pred R-Squared and PRESS statistic not defined

| Component | Coefficient |    | Standard Error | 95% CI Low | 95% CI High |
|-----------|-------------|----|----------------|------------|-------------|
|           | Estimate    | DF |                |            |             |
| A-Resin   | 41.11       | 1  | 2.23           | 36.15      | 46.07       |
| B-CX1     | -76.52      | 1  | 122.28         | -348.99    | 195.94      |
| C-CX2     | 42.32       | 1  | 4.99           | 31.21      | 53.43       |
| AB        | 191.02      | 1  | 164.41         | -175.31    | 557.35      |
| AC        | -27.84      | 1  | 14.65          | -60.47     | 4.80        |
| AD        | 9.23        | 1  | 2.23           | 4.27       | 14.19       |
| AE        | -15.48      | 1  | 2.23           | -20.44     | -10.51      |

TABLE 13.19 (Continued)

| Response: Force                                     |             |    |                |            |             |
|---|-------------|----|----------------|------------|-------------|
| ANOVA for Crossed Quadratic $\times$ 2FI Model      |             |    |                |            |             |
| Analysis of variance table [Partial sum of squares] |             |    |                |            |             |
| Component   | Coefficient |    | Standard Error | 95% CI Low | 95% CI High |
|   | Estimate    | DF |                |            |             |
| BC  | 214.55      | 1  | 162.81         | -148.22    | 577.32      |
| BD  | -87.39      | 1  | 122.28         | -359.86    | 185.07      |
| BE  | -0.92       | 1  | 122.28         | -273.38    | 271.55      |
| CD  | 6.06        | 1  | 4.99           | -5.05      | 17.17       |
| CE  | -3.20       | 1  | 4.99           | -14.31     | 7.91        |
| ABD   | 155.53      | 1  | 164.41         | -210.81    | 521.86      |
| ABE   | -26.29      | 1  | 164.41         | -392.62    | 340.04      |
| ACD   | 17.86       | 1  | 14.65          | -14.77     | 50.50       |
| ACE   | -5.49       | 1  | 14.65          | -38.12     | 27.14       |
| ADE   | -2.19       | 1  | 2.23           | -7.15      | 2.77        |
| BCD   | 165.44      | 1  | 162.81         | -197.33    | 528.21      |
| BCE   | -24.46      | 1  | 162.81         | -387.23    | 338.31      |
| BDE   | -55.00      | 1  | 122.28         | -327.47    | 217.46      |
| CDE   | 3.14        | 1  | 4.99           | -7.97      | 14.25       |
| ABDE  | 84.84       | 1  | 164.41         | -281.50    | 451.17      |
| ACDE  | -18.71      | 1  | 14.65          | -51.34     | 13.93       |
| BCDE  | 59.56       | 1  | 162.81         | -303.21    | 422.33      |

**Example 13.5 Vinyl Automobile Seat Covers** Kowalski et al. (2002) present an example of a split-plot mixture-process experiment for the production of automotive seat covers. Interest centers on producing vinyl with the desired thickness. The mixture components  $x_1$ ,  $x_2$ ,  $x_3$  are three plasticizers which can be varied from 0 to 100% of the proportion. The two process variables considered were rate of extrusion,  $z_1$ , and temperature of drying,  $z_2$ .

Based on their initial understanding of the system, the experimenters believed that a suitable model included a second-order model for the mixture variables with mixture-process and process-by-process interactions. The form of the model is

$$E(y) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \gamma_{12} z_1 z_2 \\ + \alpha_{11} x_1 z_1 + \alpha_{12} x_1 z_2 + \alpha_{21} x_2 z_1 + \alpha_{22} x_2 z_2 + \alpha_{31} x_3 z_1 + \alpha_{32} x_3 z_2 \quad (13.11)$$

Note that because of the constraint on the mixture terms, it is not possible to have linear effects for the process terms.

Since changing the rate of extrusions and temperature of drying involve reconfiguring the manufacturing environment, the experiments did not want to change this for each experimental run. Instead, the process variables were only reset seven times during the course of the experiment. Each time that the manufacturing environment had been set up, four mixtures of the plasticizers were processed, for a total of 28 runs.

Table 13.21 shows the data from the experiment, where seven different mixture combinations were used and the process variables were varied at three levels. The design structure for the process variables was a  $2^2$  factorial design with three of the whole plots selected as center runs. The chosen mixture combinations were the pure blends of each plasticizer,

TABLE 13.20 Reduced Model for Example 13.4

| Source                | Sum of Squares  | DF        | Mean Square    | F-Value      | Prob > F          |
|-----------------------|-----------------|-----------|----------------|--------------|-------------------|
| Model                 | 12009.06        | 11        | 1091.73        | 70.47        | <0.0001           |
| <i>Linear Mixture</i> | <i>640.41</i>   | <i>2</i>  | <i>320.20</i>  | <i>20.67</i> | <i>&lt;0.0001</i> |
| AC                    | 151.08          | 1         | 151.08         | 9.75         | 0.0050            |
| AD                    | 460.70          | 1         | 460.70         | 29.74        | <0.0001           |
| AE                    | 1768.65         | 1         | 1768.65        | 114.17       | <0.0001           |
| BD                    | 425.88          | 1         | 425.88         | 27.49        | <0.0001           |
| BE                    | 210.53          | 1         | 210.53         | 13.59        | 0.0013            |
| CD                    | 119.78          | 1         | 119.78         | 7.73         | 0.0109            |
| CE                    | 91.26           | 1         | 91.26          | 5.89         | 0.0239            |
| ACD                   | 61.32           | 1         | 61.32          | 3.96         | 0.0592            |
| CDE                   | 48.32           | 1         | 48.32          | 3.12         | 0.0912            |
| Residual              | 340.82          | 22        | 15.49          |              |                   |
| <i>Lack of Fit</i>    | <i>198.82</i>   | <i>17</i> | <i>11.70</i>   | <i>0.41</i>  | <i>0.9221</i>     |
| <i>Pure Error</i>     | <i>142.00</i>   | <i>5</i>  | <i>28.40</i>   |              |                   |
| <i>Cor Total</i>      | <i>12349.88</i> | <i>33</i> |                |              |                   |
| Std. Dev.             | 3.94            |           | R-Squared      |              | 0.9724            |
| Mean                  | 42.94           |           | Adj R-Squared  |              | 0.9586            |
| C.V.                  | 9.17            |           | Pred R-Squared |              | 0.9269            |
| PRESS                 | 903.20          |           | Adeq Precision |              | 28.424            |

Response: Force

ANOVA for Crossed Quadratic  $\times$  2FI Model

Analysis of variance table [Partial sum of squares]

| Component | Coefficient Estimate | DF | Standard Error | 95% CI Low | 95% CI High |
|-----------|----------------------|----|----------------|------------|-------------|
| A-Resin   | 40.66                | 1  | 1.71           | 37.11      | 44.21       |
| B-CX1     | 71.95                | 1  | 5.77           | 59.98      | 83.91       |
| C-CX2     | 46.16                | 1  | 2.68           | 40.60      | 51.71       |
| AC        | -30.58               | 1  | 9.79           | -50.88     | -10.27      |
| AD        | 9.32                 | 1  | 1.71           | 5.77       | 12.86       |
| AE        | -15.49               | 1  | 1.45           | -18.50     | -12.49      |
| BD        | 29.92                | 1  | 5.71           | 18.09      | 41.76       |
| BE        | -20.18               | 1  | 5.47           | -31.53     | -8.83       |
| CD        | 7.43                 | 1  | 2.67           | 1.89       | 12.98       |
| CE        | -4.41                | 1  | 1.82           | -8.19      | -0.64       |
| ACD       | 19.39                | 1  | 9.74           | -0.82      | 39.60       |
| CDE       | -2.60                | 1  | 1.47           | -5.65      | 0.45        |

Final Equation in Terms of Pseudo Components and Coded Factors:  
Force =

+40.66 \* A  
+71.95 \* B  
+46.16 \* C  
-30.58 \* A \* C  
+9.32 \* A \* D  
-15.49 \* A \* E  
+29.92 \* B \* D  
-20.18 \* B \* E  
+7.43 \* C \* D  
-4.41 \* C \* E  
+19.39 \* A \* C \* D  
-2.60 \* C \* D \* E

Final Equation in Terms of Actual Components:  
Force =

+73.60193 \* Resin  
-75.24237 \* CX1  
+1247.65694 \* CX2  
-1900.46726 \* Resin \* CX2  
-0.090433 \* Resin \* Temp

(continued)

TABLE 13.20 (Continued)

| Source  | Sum of Squares                     | DF              | Mean Square      | F-Value  | Prob > F | Cook's Distance | Outlier <i>t</i> | Run Order |
|---|------------------------------------|-----------------|------------------|----------|----------|-----------------|------------------|-----------|
| Response: Force                                     |                                    |                 |                  |          |          |                 |                  |           |
| ANOVA for Crossed Quadratic × 2FI Model             |                                    |                 |                  |          |          |                 |                  |           |
| Analysis of variance table [Partial sum of squares] |                                    |                 |                  |          |          |                 |                  |           |
|   |                                    |                 |                  |          |          |                 |                  |           |
|   | -0.48855 * Resin * Rel H           |                 |                  |          |          |                 |                  |           |
|   | +4.22385 * CX1 * Temp              |                 |                  |          |          |                 |                  |           |
|   | -1.16990 * CX1 * Rel H             |                 |                  |          |          |                 |                  |           |
|   | -10.55458 * CX2 * Temp             |                 |                  |          |          |                 |                  |           |
|   | +1.75779 * CX2 * Rel H             |                 |                  |          |          |                 |                  |           |
|   | + 16.2979 * Resin * CX2 * Temp     |                 |                  |          |          |                 |                  |           |
|   | -9.40951E-003 * CX2 * Temp * Rel H |                 |                  |          |          |                 |                  |           |
| Diagnostics Case Statistics                         |                                    |                 |                  |          |          |                 |                  |           |
| Standard Order                                      | Actual Value                       | Predicted Value | Student Residual | Leverage | Residual | Cook's Distance | Outlier <i>t</i> | Run Order |
| 1   | 44.00                              | 46.84           | -2.84            | 0.534    | -1.055   | 0.106           | -1.058           | 26        |
| 2   | 70.00                              | 65.47           | 4.53             | 0.534    | 1.688    | 0.272           | 1.768            | 6         |
| 3   | 19.00                              | 15.85           | 3.15             | 0.493    | 1.124    | 0.102           | 1.131            | 21        |
| 4   | 33.00                              | 34.48           | -1.48            | 0.493    | -0.528   | 0.023           | -0.520           | 8         |
| 5   | 48.00                              | 45.96           | 2.04             | 0.507    | 0.740    | 0.047           | 0.732            | 5         |
| 6   | 72.00                              | 75.97           | -3.97            | 0.425    | -1.329   | 0.109           | -1.354           | 2         |
| 7   | 32.00                              | 33.15           | -1.15            | 0.628    | -0.477   | 0.032           | -0.469           | 30        |
| 8   | 59.00                              | 55.36           | 3.64             | 0.635    | 1.534    | 0.342           | 1.585            | 24        |
| 9   | 32.00                              | 32.75           | -0.75            | 0.417    | -0.248   | 0.004           | -0.243           | 17        |
| 10  | 58.00                              | 59.72           | -1.72            | 0.355    | -0.545   | 0.014           | -0.536           | 23        |
| 11  | 21.00                              | 22.28           | -1.28            | 0.549    | -0.484   | 0.024           | -0.475           | 33        |
| 12  | 38.00                              | 41.46           | -3.46            | 0.548    | -1.306   | 0.172           | -1.329           | 9         |
| 13  | 51.00                              | 50.68           | 0.32             | 0.345    | 0.101    | 0.000           | 0.099            | 11        |
| 14  | 76.00                              | 79.61           | -3.61            | 0.380    | -1.165   | 0.069           | -1.176           | 20        |



|    |       |       |       |       |        |       |        |    |
|----|-------|-------|-------|-------|--------|-------|--------|----|
| 15 | 22.00 | 17.35 | 4.65  | 0.267 | 1.380  | 0.058 | 1.411  | 19 |
| 16 | 49.00 | 46.28 | 2.72  | 0.265 | 0.805  | 0.019 | 0.798  | 22 |
| 17 | 17.00 | 15.13 | 1.87  | 0.200 | 0.531  | 0.006 | 0.522  | 10 |
| 18 | 69.00 | 67.85 | 1.15  | 0.234 | 0.335  | 0.003 | 0.328  | 16 |
| 19 | 40.00 | 39.35 | 0.65  | 0.305 | 0.198  | 0.001 | 0.193  | 32 |
| 20 | 37.00 | 41.29 | -4.29 | 0.297 | -1.300 | 0.059 | -1.322 | 3  |
| 21 | 46.00 | 42.82 | 3.18  | 0.204 | 0.905  | 0.018 | 0.902  | 7  |
| 22 | 21.00 | 18.22 | 2.78  | 0.229 | 0.804  | 0.016 | 0.797  | 29 |
| 23 | 82.00 | 76.22 | 5.78  | 0.338 | 1.806  | 0.139 | 1.912  | 27 |
| 24 | 43.00 | 49.25 | -6.25 | 0.233 | -1.813 | 0.083 | -1.920 | 34 |
| 25 | 32.00 | 32.76 | -0.76 | 0.370 | -0.245 | 0.003 | -0.239 | 15 |
| 26 | 60.00 | 61.02 | -1.02 | 0.383 | -0.331 | 0.006 | -0.324 | 1  |
| 27 | 14.00 | 13.15 | 0.85  | 0.203 | 0.242  | 0.001 | 0.237  | 28 |
| 28 | 38.00 | 38.13 | -0.13 | 0.205 | -0.036 | 0.000 | -0.035 | 4  |
| 29 | 45.00 | 42.68 | 2.32  | 0.197 | 0.657  | 0.009 | 0.649  | 14 |
| 30 | 18.00 | 18.22 | -0.22 | 0.229 | -0.064 | 0.000 | -0.063 | 13 |
| 31 | 70.00 | 67.85 | 2.15  | 0.234 | 0.625  | 0.010 | 0.616  | 18 |
| 32 | 10.00 | 17.35 | -7.35 | 0.267 | -2.181 | 0.144 | -2.407 | 25 |
| 33 | 52.00 | 49.25 | 2.75  | 0.233 | 0.799  | 0.016 | 0.792  | 12 |
| 34 | 42.00 | 46.28 | -4.28 | 0.265 | -1.269 | 0.048 | -1.288 | 31 |

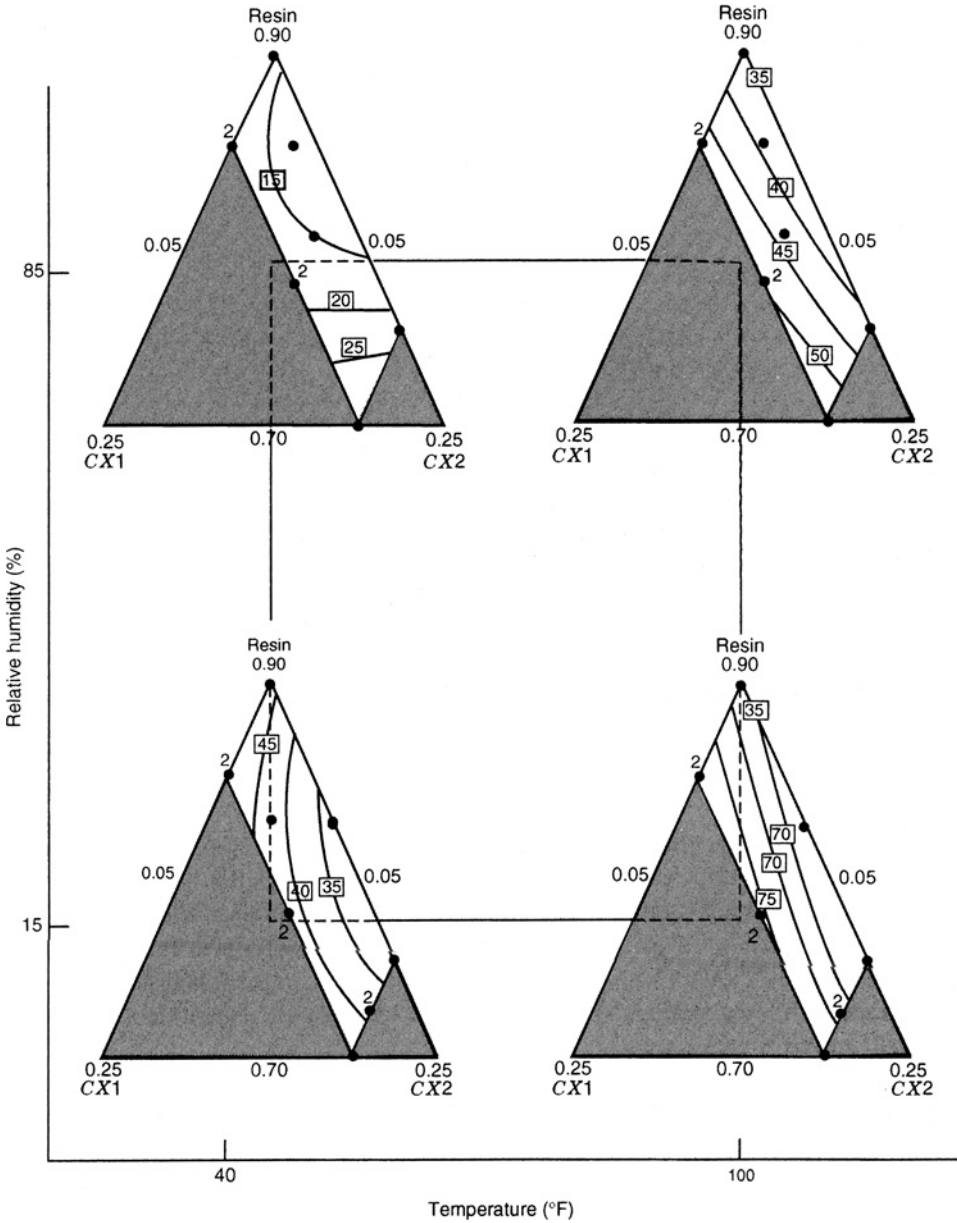
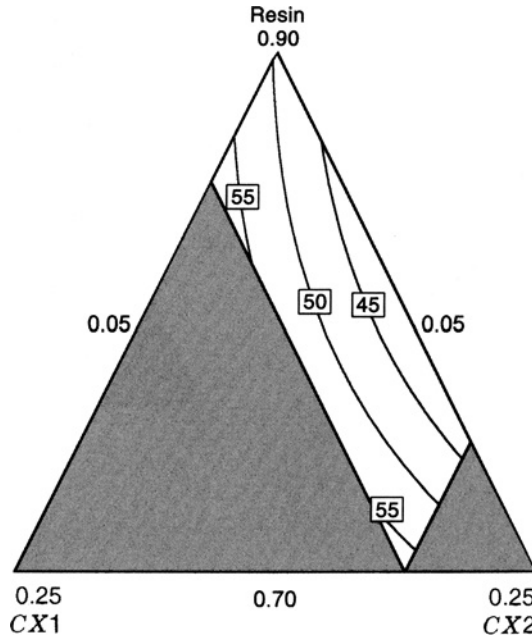


Figure 13.30 Contour plots of predicted pulloff force for Example 13.4.

two ingredient blends with equal proportions of each component, and the three component equal proportion blend.

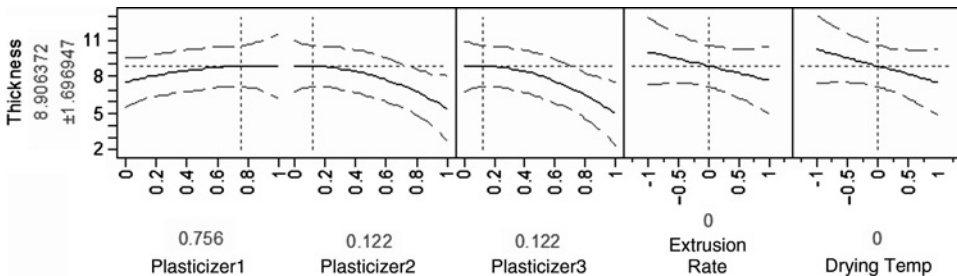
Table 13.22 shows the JMP output for this example. Because of the split plot structure, the standard error for the various terms of the model are dependent on the form of the model and estimates of the whole and subplot error terms. By using REML (Restricted or Residual Maximum Likelihood) method of fitting mixed models, we are able to estimate the two



**Figure 13.31** Contour plot of predicted pulloff force when temperature = 70°F and relative humidity = 30%.

error terms based on the data. In this case, the ratio is estimated as  $d = \sigma_{WP}^2 / \sigma^2 = 0.815$  or  $\sigma_{WP}^2 = 0.815\sigma^2$ . The interactive Prediction Profiler in JMP, shown in Fig. 13.32 allows for exploration of the design space to determine ideal combinations of plasticizers and process inputs to optimize the seat cover thickness.

The design selected for this experiment was not an ideal choice since it does not estimate the model parameters or predict in the design space as well as possible given the constraints of the experiment. Table 13.23 gives the *I*-optimal design for this experiment with the same assumed model in Equation 13.11 which significantly improves prediction throughout the design space. However, it should also be noted that the *I*-optimal design does not allocate any degrees of freedom for estimation of the pure error for the whole plot error, nor for lack of fit of the model.



**Figure 13.32** Prediction profiler for thickness of vinyl automotive seat covers.

**TABLE 13.21 Design and Data for Vinyl Automotive Seat Cover Example**

| Whole Plot | $z_1$ | $z_2$ | $x_1$ | $x_2$ | $x_3$ | Thickness |
|------------|-------|-------|-------|-------|-------|-----------|
| 1          | -1    | 1     | 1.00  | 0.00  | 0.00  | 10        |
|            |       |       | 0.00  | 1.00  | 0.00  | 8         |
|            |       |       | 0.00  | 0.00  | 1.00  | 3         |
|            |       |       | 0.33  | 0.33  | 0.33  | 8         |
| 2          | 1     | -1    | 1.00  | 0.00  | 0.00  | 10        |
|            |       |       | 0.00  | 1.00  | 0.00  | 5         |
|            |       |       | 0.00  | 0.00  | 1.00  | 9         |
|            |       |       | 0.33  | 0.33  | 0.33  | 9         |
| 3          | 1     | 1     | 0.50  | 0.50  | 0.00  | 5         |
|            |       |       | 0.50  | 0.00  | 0.50  | 4         |
|            |       |       | 0.00  | 0.50  | 0.50  | 7         |
|            |       |       | 0.33  | 0.33  | 0.33  | 10        |
| 4          | -1    | -1    | 0.50  | 0.50  | 0.00  | 7         |
|            |       |       | 0.50  | 0.00  | 0.50  | 8         |
|            |       |       | 0.00  | 0.50  | 0.50  | 4         |
|            |       |       | 0.33  | 0.33  | 0.33  | 7         |
| 5          | 0     | 0     | 0.33  | 0.33  | 0.33  | 8         |
|            |       |       | 0.33  | 0.33  | 0.33  | 7         |
|            |       |       | 0.33  | 0.33  | 0.33  | 7         |
|            |       |       | 0.33  | 0.33  | 0.33  | 8         |
| 6          | 0     | 0     | 0.33  | 0.33  | 0.33  | 7         |
|            |       |       | 0.33  | 0.33  | 0.33  | 8         |
|            |       |       | 0.33  | 0.33  | 0.33  | 9         |
|            |       |       | 0.33  | 0.33  | 0.33  | 9         |
| 7          | 0     | 0     | 0.33  | 0.33  | 0.33  | 12        |
|            |       |       | 0.33  | 0.33  | 0.33  | 10        |
|            |       |       | 0.33  | 0.33  | 0.33  | 9         |
|            |       |       | 0.33  | 0.33  | 0.33  | 11        |

### 13.3.3 Robust Parameter Designs for Mixture-Process Experiments

Often some of the process variables cannot be fully controlled during normal production, and hence we are in a robust design setting. In this case, we are interested in determining an appropriate combination of the control factors to optimize our process when the noise factors vary over their natural production ranges. As discussed in Chapter 11, it is customary to fit a response model combining control and noise variables. In this case, the control variables are likely to include both mixture and process factors. We can derive models for both the mean response and for its variance. The variance model is directly related to the directional derivative or slope in the directions of the noise variables. Adaptations of the usual robust parameter designs are needed because of the constraints introduced by the mixture components. See Steiner and Hamada (1997) and Goldfarb et al. (2003) for more details.

Similar to the development in Chapter 11, the goal of the optimization is to find a combination of the control factors that is close to the desired target for the mean, while minimizing the variability transmitted to the response through the slopes in the noise factor directions.

**TABLE 13.22 JMP Output from Data for Vinyl Automotive Seat Cover Example**

| Summary of Fit                    |           |               |           |           |           |              |
|-----------------------------------|-----------|---------------|-----------|-----------|-----------|--------------|
| RSquare                           | 0.80174   |               |           |           |           |              |
| RSquare Adj                       | 0.643132  |               |           |           |           |              |
| Root Mean Square Error            | 1.395338  |               |           |           |           |              |
| Mean of Response                  | 7.821429  |               |           |           |           |              |
| Observations (or Sum Weights)     | 28        |               |           |           |           |              |
| Parameter Estimates               |           |               |           |           |           |              |
| Term                              | Estimate  | Std Error     | DFDen     | t-Ratio   | Prob >  t |              |
| Plasticizer1(Mixture)             | 8.8940276 | 1.275453      | 14.45     | 6.97      | <.0001    |              |
| Plasticizer2(Mixture)             | 5.3940276 | 1.275453      | 14.45     | 4.23      | 0.0008    |              |
| Plasticizer3(Mixture)             | 4.8940276 | 1.275453      | 14.45     | 3.84      | 0.0017    |              |
| Extrusion Rate*Drying Temp        | -1.194491 | 0.788355      | 3.534     | -1.52     | 0.2134    |              |
| Extrusion Rate*Plasticizer1       | -2.007424 | 1.154327      | 11.79     | -1.74     | 0.1081    |              |
| Drying Temp*Plasticizer1          | -2.008637 | 1.154327      | 11.79     | -1.74     | 0.1079    |              |
| Extrusion Rate*Plasticizer2       | 0.7425761 | 1.154327      | 11.79     | 0.64      | 0.5323    |              |
| Drying Temp*Plasticizer2          | 2.2413632 | 1.154327      | 11.79     | 1.94      | 0.0765    |              |
| Extrusion Rate*Plasticizer3       | 1.9925761 | 1.154327      | 11.79     | 1.73      | 0.1104    |              |
| Drying Temp*Plasticizer3          | -1.008637 | 1.154327      | 11.79     | -0.87     | 0.3997    |              |
| Plasticizer1 *Plasticizer2        | 3.7412377 | 4.968701      | 13.83     | 0.75      | 0.4641    |              |
| Plasticizer1 *Plasticizer3        | 4.7412377 | 4.968701      | 13.83     | 0.95      | 0.3564    |              |
| Plasticizer2*Plasticizer3         | 9.7412377 | 4.968701      | 13.83     | 1.96      | 0.0704    |              |
| REML Variance Component Estimates |           |               |           |           |           |              |
| Random Effect                     | Var Ratio | Var Component | Std Error | 95% Lower | 95% Upper | Pct of Total |
| Block                             | 0.8154497 | 1.5876555     | 1.7655516 | -1.872826 | 5.0481366 | 44.917       |
| Residual                          |           | 1.9469694     | 0.7896778 | 1.004764  | 5.2623692 | 55.083       |
| Total                             |           | 3.5346249     |           |           |           | 100.000      |
| -2 LogLikelihood = 62.582878702   |           |               |           |           |           |              |

TABLE 13.23 *I*-Optimal Design for Vinyl Automotive Seat Cover Example

| Whole Plot | $z_1$ | $z_2$ | $x_1$ | $x_2$ | $x_3$ |
|------------|-------|-------|-------|-------|-------|
| 1          | 1     | 1     | 0     | 0     | 1     |
|            |       |       | 0.5   | 0     | 0.5   |
|            |       |       | 0     | 1     | 0     |
|            |       |       | 1     | 0     | 0     |
| 2          | -1    | 1     | 0     | 0     | 1     |
|            |       |       | 1     | 0     | 0     |
|            |       |       | 0     | 1     | 0     |
|            |       |       | 0.5   | 0.5   | 0     |
| 3          | -1    | -1    | 0     | 0.5   | 0.5   |
|            |       |       | 1     | 0     | 0     |
|            |       |       | 0     | 1     | 0     |
|            |       |       | 0     | 0     | 1     |
| 4          | -1    | 1     | 0     | 0.4   | 0.6   |
|            |       |       | 0     | 1     | 0     |
|            |       |       | 1     | 0     | 0     |
|            |       |       | 0.5   | 0     | 0.5   |
| 5          | -1    | -1    | 0.5   | 0.5   | 0     |
|            |       |       | 0     | 1     | 0     |
|            |       |       | 0.5   | 0     | 0.5   |
|            |       |       | 0     | 0     | 1     |
| 6          | 1     | 1     | 0     | 0     | 1     |
|            |       |       | 0     | 0.6   | 0.4   |
|            |       |       | 0.5   | 0.5   | 0     |
|            |       |       | 1     | 0     | 0     |
| 7          | 1     | -1    | 0     | 0.5   | 0.5   |
|            |       |       | 0     | 0     | 1     |
|            |       |       | 1     | 0     | 0     |
|            |       |       | 0     | 1     | 0     |

**Example 13.6 Throughput for Soap Manufacturing** Goldfarb et al. (2003) consider an example of soap manufacturing where the goal is to maximize the amount of soap produced per hour (throughput) by the process. The soap blend is comprised of three mixture components: soap ( $x_1$ ), co-surfactant ( $x_2$ ) and filler ( $x_3$ ). There are the following restrictions on the proportions of each component:

$$0.20 \leq x_1 \leq 0.80$$

$$0.15 \leq x_2 \leq 0.50$$

$$0.05 \leq x_3 \leq 0.30$$

in addition to the usual constraint that the component proportions sum to 1.

During the manufacturing process, the first process variable, mixing time ( $w_1$ ) is considered controllable, while the plodder temperature ( $z_1$ ) is difficult to control during production, and hence is considered a noise variable. Based on understanding of the production process, the chosen model for this study involved only linear mixture components and

their interactions with the controllable process and noise factors. The form of the model is

$$\begin{aligned}
 E(y) = & \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \alpha_{11}x_1w_1 + \alpha_{21}x_2w_1 + \alpha_{31}x_3w_1 \\
 & + \delta_{11}x_1z_1 + \delta_{21}x_2z_1 + \delta_{31}x_3z_1 + \lambda_{111}x_1w_1z_1 \\
 & + \lambda_{211}x_2w_1z_1 + \lambda_{311}x_3w_1z_1
 \end{aligned}
 \tag{13.12}$$

where the  $\beta$ 's are the mixture model coefficients, the  $\alpha$ 's are the interaction coefficients for the mixture and controllable process variables, the  $\delta$ 's are the interaction coefficients for the mixture and noise variables, and the  $\lambda$ 's are the interaction coefficients for the mixture, controllable process, and noise variables.

The two mixing times selected are 0.5 and 1 hour, as these cover the range common mixing times considered during production. The two plodder temperature variable levels considered are 15°C and 25°C, which is thought to cover about 70% of the variation in plodder temperature around the average value of 20°C typically observed during production. Thus the chosen values are approximately one standard deviation from the mean, and so we estimate that the mean of the noise variable is 0 and has variance  $\sigma_{z_1}^2 = 1$ .

Design-Expert was used to generate a five-run *D*-optimal design for the linear mixture model. The overall design was a completely randomized design where the generated mixture design was repeated for the four factorial combinations of the two process variables, for a total of 20 runs. Table 13.24 shows the design with the observed data.

**TABLE 13.24 Design and Date for the Throughput Soap Manufacturing Example**

| Soap<br>$x_1$ | Co-Surfactant<br>$x_2$ | Filler<br>$x_3$ | Mixing Time<br>$w_1$ | Plodder Temperature<br>$z_1$ | Throughput |
|---------------|------------------------|-----------------|----------------------|------------------------------|------------|
| 0.20          | 0.50                   | 0.30            | -1                   | -1                           | 245.2      |
| 0.80          | 0.15                   | 0.05            | -1                   | -1                           | 381.5      |
| 0.80          | 0.15                   | 0.05            | -1                   | -1                           | 381.0      |
| 0.55          | 0.15                   | 0.30            | -1                   | -1                           | 453.3      |
| 0.45          | 0.50                   | 0.05            | -1                   | -1                           | 172.9      |
| 0.20          | 0.50                   | 0.30            | 1                    | -1                           | 285.4      |
| 0.80          | 0.15                   | 0.05            | 1                    | -1                           | 409.1      |
| 0.80          | 0.15                   | 0.05            | 1                    | -1                           | 411.7      |
| 0.55          | 0.15                   | 0.30            | 1                    | -1                           | 450.0      |
| 0.45          | 0.50                   | 0.05            | 1                    | -1                           | 245.8      |
| 0.20          | 0.50                   | 0.30            | -1                   | 1                            | 213.8      |
| 0.80          | 0.15                   | 0.05            | -1                   | 1                            | 378.4      |
| 0.80          | 0.15                   | 0.05            | -1                   | 1                            | 377.3      |
| 0.55          | 0.15                   | 0.30            | -1                   | 1                            | 408.8      |
| 0.45          | 0.50                   | 0.05            | -1                   | 1                            | 180.6      |
| 0.20          | 0.50                   | 0.30            | 1                    | 1                            | 250.6      |
| 0.80          | 0.15                   | 0.05            | 1                    | 1                            | 404.2      |
| 0.80          | 0.15                   | 0.05            | 1                    | 1                            | 406.6      |
| 0.55          | 0.15                   | 0.30            | 1                    | 1                            | 410.1      |
| 0.45          | 0.50                   | 0.05            | 1                    | 1                            | 245.4      |

The fitted model was reduced by eliminating  $\lambda_{111}x_1w_1z_1$ , the only nonsignificant term (using a cutoff of 0.05) from Equation 13.12 to obtain the final model:

$$\hat{y} = 438.36x_1 - 254.01x_2 + 765.78x_3 + 34.55x_1w_1 + 263.49x_2w_1 - 203.75x_3w_1 - 0.04x_1z_1 + 5.27x_2z_1 - 18.56x_3z_1 - 3.90x_2w_1z_1 + 4.69x_3w_1z_1 \quad (13.13)$$

We can estimate the mean response surface for throughput as a function of the mixture and controllable process variables as

$$\hat{E}(\hat{y}) = 438.36x_1 - 254.01x_2 + 765.78x_3 + 34.55x_1w_1 + 263.49x_2w_1 - 203.75x_3w_1$$

since the mean of the noise variable is assumed to be zero in the coded variables. The estimated response surface model for the throughput variance is calculated as the derivative of the overall fitted model in Equation 13.13 with respect to the noise variable. In this case it is estimated to be

$$\text{Var}(y) = \hat{\sigma}_{z_1}^2 (-0.04x_1 + 5.27x_2 - 18.56x_3 - 3.90x_2w_1 + 4.69x_3w_1)^2 + \hat{\sigma}^2$$

with an estimate of  $\hat{\sigma}^2 = 1.07$ , using the mean squared error for the fitted model.

To optimize the process, we choose a combination of the mixture components and the mixing time which simultaneously maximizes the average throughput, while also minimizing the variability. Using the desirability function optimizer in Design-Expert, we obtain the following optimal conditions for production:

$$\text{Soap } (x_1) = 0.80$$

$$\text{Co-surfactant } (x_2) = 0.15$$

$$\text{Filler } (x_3) = 0.05$$

$$\text{Mixing Time } (w_1) = 1 \text{ hour (+ 1 in coded variables).}$$

The predicted throughput at these factor settings is 408.5 with a standard deviation of 0.52.

Since robust parameter design involves simultaneously optimizing the mean and variance of the process, different tools are needed to assess the goodness of the design for this dual purpose. Goldfarb et al. (2004c) developed variations of the variance dispersion graphs and fraction of designs space plots to evaluate and compare different designs based on both of these criteria. Goldfarb et al. (2005) and Chung et al. (2007) used genetic algorithms to generate designs which perform well for both the mean and slope models. By using a desirability function to combine the prediction variance for both mean and slope, optimal designs for robust parameter design experiments can be created. Similar to the robust parameter design settings with just continuous process factors, designs that only consider the mean model may have very poor performance for the slope. However, with only small modifications, a design which is nearly optimal for the mean can often be adapted to perform well for both the mean and slope prediction variances.



### 13.4 SCREENING MIXTURE COMPONENTS

In some mixture experiments there are a large number of components, and the initial objective of the experimenter is to screen these components to identify the ones that are most important. In fact, any time there are six or more components ( $q \geq 6$ ) the experimenter should consider a **screening experiment** to reduce the number of components.

Screening is often done using the first-order mixture model

$$E(y) = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_q x_q \quad (13.14)$$

If the experimental region is a simplex, it is usually a good idea to make the ranges of the components as similar as possible, because then the relative effects of the components can be assessed by ranking the ratios of the parameter estimates  $\beta_i$ ,  $i = 1, 2, \dots, q$ , relative to their standard errors. Obviously, important components will have large values of  $b_i/se(b_i)$ . If the effects of two or more components are equal, then the proportions of the individual components are summed and their sum is considered to be a new component.

An **axial design** in which the design points are placed only on the component axes is often recommended for screening mixture components. The simplest type of axial design has points positioned equidistant from the centroid toward each vertex. Generally, as the number of components increases, the spread of these axial points should be increased in order to increase the precision of the parameter estimates. If the vertices of the simplex are complete mixtures, as will be the case if only upper-bound constraints are active, we recommend placing the axial runs at the vertices. On the other hand, if the vertices are pure blends, some experimenters prefer to run the axial check blends as the axial design, because every run in the design will be a complete mixture.

It is, of course, still possible to examine the ratios  $b_i/se(b_i)$  in situations where the constraints form a region of experimentation that is not a simplex. We usually suggest a  $D$ -optimal design for the linear model (Eq. 13.14) in these cases.

To screen out unimportant components, it is necessary to measure the **effects** of the individual components. In general, we define the effect of component  $i$  as the change in the expected response from a change in the proportion of component  $i$  while holding the relative proportions of the other components constant. The largest change that we can make in component  $i$  is from 0 to 1. To keep the proportions of the other components constant, this change must be made along the  $x_i$ -axis. The estimate of the response when  $x_i = 0$  is

$$\begin{aligned} \hat{y} &= \sum_{j \neq i}^q b_j \left( \frac{1}{q-1} \right) \\ &= (q-1)^{-1} \sum_{j \neq i}^q b_j \end{aligned}$$

Similarly, at the vertex where  $x_i = 1$ , the predicted response is

$$\hat{y} = b_i$$

If the first-order model is adequate, the **effect of component  $i$**  is just the difference in these two predicted responses:

$$(\text{effect of component } i) = b_i - (q-1)^{-1} \sum_{j \neq i}^q b_j, \quad i = 1, 2, \dots, q \quad (13.15)$$

This equation can be used for model reduction. If the true surface is linear and

$$b_i = (q - 1)^{-1} \sum_{j \neq i}^q b_j$$

then the effect of component  $i$  is zero. Then the term  $\beta_i x_i$  can be removed from the model, and the component  $x_i$  is not considered further in the experiment.

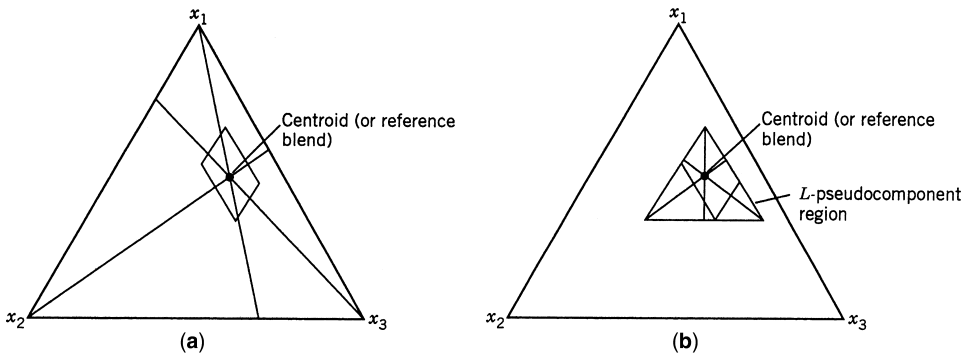
When the feasible region is constrained, then often we find that the ranges of the individual components are not equal. This renders the estimate of an effect in Equation 13.15 less useful. A simple way to remedy this situation is to **adjust** the component effect estimate as follows:

$$\text{adjusted effect of component } i = R_i \left[ b_i - (q - 1)^{-1} \sum_{j \neq i}^q b_j \right], \quad (13.16)$$

$$i = 1, 2, \dots, q$$

where  $R_i = U_i - L_i$  is the range of component  $i$ . Notice that the adjustment essentially weights the component effect by the range of the  $i$ th component. Piepel (1982) and Cornell (2002) also discuss measuring component effects in constrained regions.

The response trace plots introduced in Chapter 12 (Section 12.3) are also very useful in screening mixture components. If one (or more) of the curves on the response trace plot is a horizontal line, then we have detected an **inactive ingredient**. In constrained regions that are not simplexes, care must be taken in defining both the reference mixture and the direction for measuring the component effect. Generally, the centroid of the constrained region is chosen as the reference mixture, and the effect of the  $i$ th component is measured along a line connecting this centroid to the vertex  $x_i = 1$ . This direction was introduced by Cox (1971) in presenting new mixture models as alternatives to those proposed by Scheffé. Consequently, this is sometimes called **Cox's direction**. The construction of the response surface trace plot for constrained regions is then identical to the procedure discussed in Section 12.3 for a simplex region of experimentation. The trace plot in Fig. 13.23 (the shampoo foam experiment, Example 13.2) uses Cox's direction. Adjusting the proportion of



**Figure 13.33** Illustration of (a) Cox's direction and (b) Piepel's direction for the shampoo foam experiment, Example 13.1.

one component along Cox's direction results in the ratio of the other components remaining constant. Piepel (1982) introduced another direction that could be used to construct trace plots. His direction is the line extending from a reference mixture (usually the centroid of the constrained region) to the vertices of the  $L$ -pseudocomponent region. Figure 13.33 illustrates and compares these two directions for the shampoo foam experiment. These two directions are closely related, and usually there is little difference in the trace plots.

## EXERCISES

- 13.1** Suppose that you have fitted a second-order mixture model using  $L$ -pseudocomponents for  $q = 3$ . The fitted model is

$$\hat{y} = b_1X_1 + b_2X_2 + b_3X_3 + b_{12}X_1X_2 + b_{13}X_1X_3 + b_{23}X_2X_3$$

You wish to convert this to a model in the original components, say

$$\hat{y} = a_1x_1 + a_2x_2 + a_3x_3 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3$$

Show that the  $a$ 's can be expressed in the terms of the  $b$ 's as follows:

$$a_1 = \frac{b_{12}L_2(L_1 - 1) + b_{13}L_3(L_1 - 1) + b_{23}L_2L_3}{(1 - L)^2} + \frac{b_1 - \sum_{i=1}^3 b_iL_i}{1 - L}$$

$$a_2 = \frac{b_{12}L_1(L_2 - 1) + b_{13}L_1L_3 + b_{23}L_3(L_2 - 1)}{(1 - L)^2} + \frac{b_2 - \sum_{i=1}^3 b_iL_i}{1 - L}$$

$$a_3 = \frac{b_{12}L_1L_2 + b_{13}L_1(L_3 - 1) + b_{23}L_2(L_3 - 1)}{(1 - L)^2} + \frac{b_3 - \sum_{i=1}^3 b_iL_i}{1 - L}$$

$$a_{ij} = b_{ij}/(1 - L)^2, \quad i, j = 1, 2, 3, \quad i < j$$

- 13.2** An article in *Industrial Quality Control* ("Experiments with Mixtures of Components Having Lower Bounds," by I. S. Kurotori, Vol. 22, 1966, pp. 592–596) describes an experiment to investigate the formulation of a rocket propellant. The propellant consists of fuel ( $x_1$ ), oxidizer ( $x_2$ ), and binder ( $x_3$ ). The restrictions on the component proportions are  $0.20 \leq x_1$ ,  $0.40 \leq x_2$ , and  $0.20 \leq x_1 \leq 0.40$ , with  $x_1 + x_2 + x_3 = 1$ . The response variable of interest is elasticity.

- Draw a graph of the feasible region.
- The data collected by Kurotori are shown in Table E13.1.  
What type of experimental design has been used? Critique the choice of design.
- Fit an appropriate model relating elasticity to the component proportions (use the pseudocomponents).
- Convert the model found in part (c) to an equivalent model in the original component proportions.
- Plot the contours of constant elasticity, and interpret the fitted response surface.

**TABLE E13.1 Data for Exercise 13.2**

| Design Point | $x_1$ | $x_2$ | $x_3$ | $X_1$ | $X_2$ | $X_3$ | Elasticity |
|--------------|-------|-------|-------|-------|-------|-------|------------|
| 1            | 0.40  | 0.40  | 0.40  | 1     | 0     | 0     | 2650       |
| 2            | 0.20  | 0.60  | 0.20  | 0     | 1     | 0     | 2450       |
| 3            | 0.20  | 0.40  | 0.40  | 0     | 0     | 1     | 2350       |
| 4            | 0.30  | 0.50  | 0.20  | 0.50  | 0.50  | 0     | 2950       |
| 5            | 0.30  | 0.40  | 0.30  | 0.50  | 0     | 0.50  | 2750       |
| 6            | 0.20  | 0.50  | 0.30  | 0     | 0.50  | 0.50  | 2400       |
| 7            | 0.27  | 0.46  | 0.27  | 0.333 | 0.33  | 0.333 | 3000       |
| 8            | 0.33  | 0.43  | 0.23  | 0.667 | 0.166 | 0.166 | 2980       |
| 9            | 0.23  | 0.53  | 0.23  | 0.166 | 0.667 | 0.166 | 2770       |
| 10           | 0.23  | 0.43  | 0.33  | 0.166 | 0.166 | 0.667 | 2690       |

- 13.3** Koons and Wilt (1985) describe an experiment involving the formulation of acrylonitrile–butadiene styrene (ABS), a material used to make plastic pipe. ABS is normally made from two materials: grafted polybutadiene ( $x_1 = \text{graft}$ ) and styrene–acrylonitrile ( $x_2 = \text{SAN}$ ). Mixture experiments were performed to determine whether or not coal-tar pitch ( $x_3 = \text{pitch}$ ) could be added to the SAN–graft combination and form a pipe that would have physical properties that met ASTM specifications. The motive was to reduce pipe cost and utilize a by-product produced elsewhere in the plant. Experience indicated that the following constraints (on weight) should be used:

$$0.45 < x_1 < 0.70$$

$$0.30 < x_2 < 0.55$$

$$0.0 < x_3 < 0.25$$

The physical properties of interest for the ABS pipe are: Izod impact strength ( $y_1$ , ft-lb/in), deflection temperature under load ( $y_2$ , °F), and yield strength ( $y_3$ , psi).

- (a) Graph the feasible region for this experiment.  
 (b) The response data obtained in the experiment performed by Koons and Wilt are shown in Table E13.2.

**TABLE E13.2 Data for Exercise 13.3**

| Run | $x_1$ | $x_2$ | $x_3$ | $X_1$ | $X_2$ | $X_3$ | $y_1$ | $y_2$ | $y_3$ |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1   | 0.700 | 0.300 | 0.000 | 1     | 0     | 0     | 7.4   | 205   | 4350  |
| 2   | 0.450 | 0.550 | 0.000 | 0     | 1     | 0     | 6.1   | 213   | 6080  |
| 3   | 0.450 | 0.300 | 0.250 | 0     | 0     | 1     | 0.8   | 183   | 5065  |
| 4   | 0.575 | 0.425 | 0.000 | 0.50  | 0.50  | 0     | 7.3   | 204   | 5280  |
| 5   | 0.575 | 0.300 | 0.125 | 0.50  | 0     | 0.50  | 3.9   | 188   | 5080  |
| 6   | 0.450 | 0.425 | 0.125 | 0     | 0.50  | 0.50  | 2.1   | 191   | 5740  |
| 7   | 0.535 | 0.380 | 0.085 | 0.34  | 0.32  | 0.34  | 4.4   | 197   | 5410  |
| 8   | 0.600 | 0.350 | 0.050 | 0.60  | 0.20  | 0.20  | 6.1   | 192   | 5035  |
| 9   | 0.500 | 0.450 | 0.050 | 0.20  | 0.60  | 0.20  | 4.9   | 202   | 5615  |
| 10  | 0.500 | 0.350 | 0.150 | 0.20  | 0.20  | 0.60  | 2.5   | 186   | 5385  |

- (c) What type of experimental design has been used? Comment on the experimenter's choice of design.
- (d) Build appropriate response surface models for all three responses. Build these models using the pseudocomponents.
- (e) Convert the models found in part (d) to equivalent models in the actual component proportions.
- (f) Construct response surface contour plots for the three response variables  $y_1$ ,  $y_2$ , and  $y_3$ .
- (g) In order to produce acceptable pipe from this modified ABS material, the following response constraints must be satisfied:  $y_1 > 1$  ft-lb/in.,  $y_2 > 190^\circ\text{F}$ , and  $y_3 > 5000$  psi. Is there a formulation of the modified ABS material that will satisfy these specifications?
- (h) If the objective is to include as much pitch in the formulation as possible and still satisfy the response specifications in part (g), what formulation would you recommend?

**13.4** The following constraints are imposed on three mixture components:

$$\begin{aligned} 0.15 \leq x_1, & & 0.25 \leq x_2, & & 0.10 \leq x_3 \\ & & x_1 + x_2 + x_3 = 1 & & \end{aligned}$$

- (a) Draw a graph of the feasible region.
  - (b) Suppose that you could only afford to perform eight runs, and your objective is to fit a first-order mixture model. What runs would you recommend?
  - (c) Suppose that your objective is to fit a quadratic mixture model. If you can afford to perform 12 runs, what blends would you include in the design?
  - (d) List the points in the design from part (c) in terms of the  $L$ -pseudocomponents.
- 13.5** Reconsider the three component mixture problem in Exercise 13.4. What design would you recommend if the experimenter wishes to fit a special cubic model?
- 13.6** Table E13.3 presents the actual values of the observed response from the shampoo foam height experiment in Example 13.2, the predicted response from the quadratic model, the residuals, the values of  $h_{ii}$ , and the studentized residuals.
- (a) Construct a normal probability plot of the studentized residuals from this experiment. Interpret the plot.
  - (b) Plot the studentized residuals versus the predicted values of foam height. What conclusions can you show about model adequacy?
  - (c) Two points in the design (a) have relative high leverage ( $h_{10,10} = 0.874$ ,  $h_{11,11} = 0.909$ ). Where are these points located in the feasible space? Discuss the practical implications of having design points with high average.
- 13.7** A three-component mixture experiment is subject to the following constraints on the component proportions:

$$\begin{aligned} 0.3 \leq x_1 \leq 0.6, & & 0.1 \leq x_2 \leq 0.5, & & 0.05 \leq x_3 \leq 0.45 \\ & & x_1 + x_2 + x_3 = 1 & & \end{aligned}$$

- (a) Graph the feasible region for this experiment.

**TABLE E13.3 Data for Exercise 13.6**

| Run | Actual Value | Predicted Value | Residual | $h_{ii}$ | Studentized Residual |
|-----|--------------|-----------------|----------|----------|----------------------|
| 1   | 152.00       | 146.74          | 5.26     | 0.484    | 1.461                |
| 2   | 140.00       | 146.74          | -6.74    | 0.484    | -1.872               |
| 3   | 150.00       | 148.16          | 1.84     | 0.466    | 0.501                |
| 4   | 145.00       | 148.16          | -3.16    | 0.466    | -0.863               |
| 5   | 141.00       | 139.43          | 1.57     | 0.499    | 0.443                |
| 6   | 138.00       | 139.43          | -1.43    | 0.499    | -0.403               |
| 7   | 153.00       | 148.93          | 4.07     | 0.473    | 1.119                |
| 8   | 147.00       | 148.93          | -1.93    | 0.473    | -0.529               |
| 9   | 165.00       | 164.22          | 0.78     | 0.617    | 0.250                |
| 10  | 170.00       | 170.28          | -0.28    | 0.874    | -0.156               |
| 11  | 148.00       | 146.95          | 1.05     | 0.909    | 0.698                |
| 12  | 175.00       | 171.20          | 3.80     | 0.346    | 0.936                |
| 13  | 163.00       | 167.83          | -4.83    | 0.409    | -1.254               |

- (b) Suppose that the experimenter wishes to fit a quadratic model to the response. Identify a set of candidate points that the experimenter should consider in the selected design.
- (c) Find a  $D$ -optimal design for this experiment with  $n = 14$  runs, assuming that four of these runs must be replicates.
- (d) Find an  $I$ -optimal design for this experiment with  $n = 14$  runs.
- (e) Construct an FDS plot for each of the designs in parts (c) and (d)
- 13.8** Reconsider the constrained mixture problem in Exercise 13.7.
- (a) Find a 10-point design for this experiment using the distance criterion, assuming that you are not required to replicate any points.
- (b) Find a 14-point distance-criterion design for this experiment, assuming that four of these points must be replicates.
- 13.9** Reconsider the three-component mixture experiment from Exercise 13.7. Suppose the experimenter wishes to fit a linear mixture model to the response and is willing to make  $n = 8$  runs.
- (a) Construct a  $D$ -optimal design for this problem.
- (b) Construct a distance-criterion design for this problem.
- (c) Construct an FDS plot for each design in parts (a) and (b)
- (d) Compare and contrast the designs found in parts (a) and (b). Which design would you recommend?
- 13.10** Suppose that the experimenter in Exercise 13.7 is interested in fitting a special cubic model.
- (a) Construct an FDS plot for the designs.
- (b) Compare the prediction variance for each design using the linear model, the quadratic model, and the special cubic.
- (c) Which design would you recommend?

- 13.11** A four-component mixture experiment is subject to the following constraints on the component proportions:

$$0.8 \leq x_1, \quad 0.05 \leq x_2 \leq 0.15, \quad 0.02 \leq x_3 \leq 0.10, \quad x_4 \leq 0.05$$

$$x_1 + x_2 + x_3 + x_4 = 1$$

- (a) Suppose that the experimenter wishes to fit a quadratic model to the response. Identify a set of candidate points that the experimenter should consider in the selected design.
- (b) Find a  $D$ -optimal design for this experiment with  $n = 15$  runs, assuming that you are not required to replicate any points.
- (c) Find a  $D$ -optimal design for this experiment with  $n = 18$  runs, assuming that four of these runs must be replicates.
- (d) Find  $I$ -optimal designs for this experiment with  $n = 15$  and  $n = 18$  runs.
- (e) Construct an FDS plot for the four designs in parts (b), (c), and (d).
- (f) Which design would you recommend?
- 13.12** Consider the constrained mixture problem in Exercise 13.11.
- (a) Find a 15-point design for this experiment using the distance criterion, assuming that you are not required to replicate any points.
- (b) Find an 18-point design for this experiment using the distance criterion. Assume that four runs must be replicates.
- (c) Construct an FDS plot for the designs in parts (a) and (b).
- (d) Which design would you recommend?
- 13.13** Consider the four-component mixture experiment from Exercise 13.11. Suppose that the experimenter wishes to fit a linear mixture model to the response and is willing to make  $n = 10$  runs.
- (a) Construct a  $D$ -optimal design for this problem.
- (b) Construct a distance-based design for this problem.
- (c) Compare and contrast the designs in parts (a) and (b). Which design would you recommend?
- 13.14** Heinsman and Montgomery (1995) describe a four-component mixture experiment used to optimize the formulation of a household detergent. The constraints on the component proportions are:

$$0.5 \leq x_1 \leq 1, \quad 0 \leq x_2 \leq 0.5, \quad 0 \leq x_3 \leq 0.5, \quad 0 \leq x_4 \leq 0.05$$

$$x_1 + x_2 + x_3 + x_4 = 1$$

The authors measured four responses:  $y_1$  = product life (in lather units),  $y_2$  = soil pellets,  $y_3$  = foam height, and  $y_4$  = total foam. The last three responses reflect the grease-cutting ability of the product. All four responses should be as large as possible. The data from the experiment is shown in Table E13.4. The design used by the authors was generated by selecting some points with the distance criterion and the remaining points with the  $D$ -optimal criterion.

- (a) Build response surface models for all four responses.
- (b) Plot contours of each response, and interpret the fitted surfaces.
- (c) What formulation of these ingredients would you recommend?

TABLE E13.4 Experimental Design for Exercise 13.14

| Observation | $x_1$ | $x_2$  | $x_3$  | $x_4$   | $X_1$ | $X_2$ | $X_3$ | $X_4$  | $y_1$ (Lather Units) | $y_2$ (Pellets) | $y_3$ (Foam Height) | $y_4$ (Total Foam) |
|-------------|-------|--------|--------|---------|-------|-------|-------|--------|----------------------|-----------------|---------------------|--------------------|
| 1           | 1.000 | 0.0000 | 0.0000 | 0.00000 | 1.000 | 0.000 | 0.000 | 0.0000 | 7.170                | 7.00            | 95.0                | 559.0              |
| 2           | 0.500 | 0.5000 | 0.0000 | 0.00000 | 0.000 | 1.000 | 0.000 | 0.0000 | 2.680                | 20.00           | 92.0                | 1320.0             |
| 3           | 0.500 | 0.0000 | 0.5000 | 0.00000 | 0.000 | 0.000 | 1.000 | 0.0000 | 3.080                | 3.00            | 44.0                | 275.0              |
| 4           | 0.950 | 0.0000 | 0.0000 | 0.05000 | 0.900 | 0.000 | 0.000 | 0.1000 | 6.990                | 7.00            | 73.0                | 508.0              |
| 5           | 0.500 | 0.4500 | 0.0000 | 0.05000 | 0.000 | 0.900 | 0.000 | 0.1000 | 2.920                | 20.00           | 105.0               | 1436.0             |
| 6           | 0.500 | 0.0000 | 0.4500 | 0.05000 | 0.000 | 0.000 | 0.900 | 0.1000 | 2.890                | 5.00            | 45.0                | 371.0              |
| 7           | 0.750 | 0.2500 | 0.0000 | 0.00000 | 0.500 | 0.500 | 0.000 | 0.0000 | 4.830                | 20.00           | 88.0                | 12.1               |
| 8           | 0.750 | 0.0000 | 0.2500 | 0.00000 | 0.500 | 0.000 | 0.500 | 0.0000 | 3.850                | 8.00            | 53.0                | 510.0              |
| 9           | 0.500 | 0.2500 | 0.2500 | 0.00000 | 0.000 | 0.500 | 0.500 | 0.0000 | 3.130                | 20.00           | 70.0                | 1123.0             |
| 10          | 0.725 | 0.2250 | 0.0000 | 0.05000 | 0.450 | 0.450 | 0.000 | 0.1000 | 4.430                | 20.00           | 80.0                | 1196.0             |
| 11          | 0.725 | 0.0000 | 0.2250 | 0.05000 | 0.450 | 0.000 | 0.450 | 0.1000 | 3.600                | 8.00            | 75.0                | 581.0              |
| 12          | 0.650 | 0.1500 | 0.1500 | 0.05000 | 0.300 | 0.300 | 0.300 | 0.1000 | 3.750                | 20.00           | 58.0                | 1061.0             |
| 13          | 0.650 | 0.1500 | 0.1500 | 0.05000 | 0.300 | 0.300 | 0.300 | 0.1000 | 3.260                | 20.00           | 59.0                | 1087.0             |
| 14          | 0.829 | 0.0792 | 0.0792 | 0.01250 | 0.658 | 0.158 | 0.158 | 0.0250 | 5.390                | 8.00            | 65.0                | 546.0              |
| 15          | 0.658 | 0.1583 | 0.1583 | 0.02500 | 0.317 | 0.317 | 0.317 | 0.0500 | 4.310                | 20.00           | 55.0                | 1069.0             |
| 16          | 0.579 | 0.3292 | 0.0792 | 0.01250 | 0.158 | 0.658 | 0.158 | 0.0250 | 2.640                | 20.00           | 80.0                | 1310.0             |
| 17          | 0.579 | 0.0792 | 0.3292 | 0.01250 | 0.158 | 0.158 | 0.658 | 0.0250 | 3.560                | 20.00           | 57.0                | 1011.0             |
| 18          | 0.804 | 0.0792 | 0.0792 | 0.03750 | 0.608 | 0.158 | 0.158 | 0.0750 | 5.230                | 20.00           | 68.0                | 1039.0             |
| 19          | 0.579 | 0.3042 | 0.0792 | 0.03750 | 0.158 | 0.608 | 0.158 | 0.0750 | 3.220                | 20.00           | 76.0                | 1192.0             |
| 20          | 0.579 | 0.0792 | 0.3042 | 0.03750 | 0.158 | 0.158 | 0.608 | 0.0750 | 3.520                | 20.00           | 59.0                | 1087.0             |



**TABLE E13.5 Fish Patty Texture for Exercises 13.15 and 13.16**

| Process Factors |       |       | Mixture Composition |           |           |                                 |                                 |                                 |   |
|-----------------|-------|-------|---------------------|-----------|-----------|---------------------------------|---------------------------------|---------------------------------|---|
| $z_1$           | $z_2$ | $z_3$ | (1, 0, 0)           | (0, 1, 0) | (0, 0, 1) | $(\frac{1}{2}, \frac{1}{2}, 0)$ | $(\frac{1}{2}, 0, \frac{1}{2})$ | $(0, \frac{1}{2}, \frac{1}{2})$ | $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ |
| -1              | -1    | -1    | 1.84                | 0.67      | 1.51      | 1.29                            | 1.42                            | 1.16                            | 1.59                                      |
| 1               | -1    | -1    | 2.86                | 1.10      | 1.60      | 1.53                            | 1.81                            | 1.50                            | 1.68                                      |
| -1              | 1     | -1    | 3.01                | 1.21      | 2.32      | 1.93                            | 2.57                            | 1.83                            | 1.94                                      |
| 1               | 1     | -1    | 4.13                | 1.67      | 2.57      | 2.26                            | 3.15                            | 2.22                            | 2.60                                      |
| -1              | -1    | 1     | 1.65                | 0.58      | 1.21      | 1.18                            | 1.45                            | 1.07                            | 1.41                                      |
| 1               | -1    | 1     | 2.32                | 0.97      | 2.12      | 1.45                            | 1.93                            | 1.28                            | 1.54                                      |
| -1              | 1     | 1     | 3.04                | 1.16      | 2.00      | 1.85                            | 2.39                            | 1.60                            | 2.05                                      |
| 1               | 1     | 1     | 4.13                | 1.30      | 2.75      | 2.06                            | 2.82                            | 2.10                            | 2.32                                      |

**13.15 John Cornell's Famous Fish Patties I.** Cornell (2002) describes an experiment to optimize the texture of fish patties made from three types of fish: mullet ( $x_1$ ), sheepshead ( $x_2$ ), and croaker ( $x_3$ ). The texture is measured by the amount of force required to break the patty. Three process variables are also considered in the experiment: oven temperature (375°F, 425°F), time in oven (25 sec, 40 min), and deep fat frying time (25 sec, 40 sec). The results of this experiment are shown in Table E13.5, where  $z_1$  = oven temperature,  $z_2$  = time in oven, and  $z_3$  = deep fat frying time shown in the usual  $-1, +1$  coded units. In this problem we will consider only one of the process variables,  $z_1$  = oven temperature.

- Fit a special cubic model to the texture response at the high temperature level only. Construct a contour plot of the response surface, and interpret the results.
- Fit a special cubic model to the texture response at the low temperature level only. Construct a contour plot of the response surface, and interpret the results.
- Fit a combined model to the texture response that is a special cubic in the mixture components. Is there a simple relationship between the parameter estimates in the combined model and the parameter estimates in the two individual models built at the low and high temperature levels? Does this relationship seem reasonable?
- What process conditions and recipe produces texture readings above 2.5?

**13.16 John Cornell's Famous Fish Patties II.** Reconsider the fish patty experiment described in Exercise 13.15.

- Fit a combined model to the texture response data. Can this model be **reduced** by eliminating apparently nonsignificant variables?
- Construct contour plots that explain the results of your model-building efforts.
- Suppose that it is desirable to have the texture above 2.5 and as close to 2.75 as possible. Find a recipe and a set of processing conditions that all satisfy these requirements.

**13.17** An experiment was conducted to study the hardness and percentage of solids in an automotive clear coat. There are two categorical process factors: monomer type (M1 and M2), and crosslinker type (X1, X2, X3). The experimenters want to examine

the effect of varying the proportions of monomer, crosslinker, and resin in a mixture process variables experiment. The constraints on the mixture components are:

|             |        |
|-------------|--------|
| Monomer     | 5–20%  |
| Crosslinker | 25–40% |
| Resin       | 55–70% |

A  $D$ -optimal design with 38 runs was constructed for this experiment. The experimenters assumed that only main effects of the process variables were likely to be important, but they planned to use (up to) a special cubic model in the mixture components. The design is shown in Table E13.6.

- How would you create the set of candidate points from which to construct the  $D$ -optimal design?
  - Analyze the data and select appropriate models for both responses: hardness and percentage of solids.
  - Construct contour plots of both responses.
  - Can you find a recipe for the clear coat, a monomer type, and a crosslinker type for which the hardness exceeds 10 and solids exceed 50%?
  - Can you specify the mixture recipe so that the responses are robust to the type of crosslinker?
  - Can you specify the mixture recipe so that the responses are robust to the type of monomer?
- 13.18** Reconsider John Cornell's famous fish patty experiment described in Exercise 13.15. Suppose that only two of the process variables are considered important:  $z_1$  and  $z_2$ . However, it is considered necessary to fit a model that is quadratic in the process variables.
- Write down an appropriate crossed model for this problem.
  - Set up a  $D$ -optimal design that could be used to fit the model.
- 13.19** Consider a three-component mixture experiment with the following constraints on the mixture proportions:

$$x_1 \leq 0.5, \quad x_2 \leq 0.6, \quad x_3 \leq 0.6$$

$$x_1 + x_2 + x_3 = 1$$

- Graph the feasible region for this experiment.
  - Suppose that the experimenter considers a quadratic model to be adequate. What type of experimental design would you recommend? Construct the design.
- 13.20** Consider the following first-order mixture model in seven components:

$$\hat{y} = 35x_1 + 85x_2 + 140x_3 + 77x_4 + 90x_5 + 100x_6 + 120x_7$$

- Assuming that all components were varied over approximately the same ranges, calculate the effects of each component.
- Is there any indication that some components can be dropped from further consideration?

**TABLE E13.6 Data for Exercise 13.17**

| Std | Run | Block | Component 1<br>Monomer (%) | Component 2<br>Crosslinker (%) | Component 3,<br>Resin (%) | Factor 1,<br>Vendor Type | Factor 2,<br>Crosslinker Type | Response 1,<br>Hardness | Response 2,<br>Solids (%) |
|-----|-----|-------|----------------------------|--------------------------------|---------------------------|--------------------------|-------------------------------|-------------------------|---------------------------|
| 1   | 26  | 1     | 0.000                      | 0.000                          | 1.000                     | M2                       | X3                            | 8.6                     | 50.7                      |
| 2   | 35  | 1     | 0.000                      | 1.000                          | 0.000                     | M1                       | X1                            | 9.7                     | 48.0                      |
| 3   | 23  | 1     | 0.500                      | 0.500                          | 0.000                     | M1                       | X1                            | 15.0                    | 48.3                      |
| 4   | 24  | 1     | 0.333                      | 0.333                          | 0.333                     | M2                       | X2                            | 6.4                     | 50.7                      |
| 5   | 19  | 1     | 0.000                      | 0.000                          | 1.000                     | M1                       | X1                            | 7.5                     | 49.7                      |
| 6   | 27  | 1     | 1.000                      | 0.000                          | 0.000                     | M2                       | X1                            | 12.0                    | 47.3                      |
| 7   | 4   | 1     | 1.000                      | 0.000                          | 0.000                     | M1                       | X2                            | 3.8                     | 48.0                      |
| 8   | 32  | 1     | 0.333                      | 0.333                          | 0.333                     | M1                       | X1                            | 11.0                    | 52.0                      |
| 9   | 18  | 1     | 0.000                      | 1.000                          | 0.000                     | M2                       | X1                            | 9.3                     | 48.1                      |
| 10  | 17  | 1     | 0.000                      | 0.000                          | 1.000                     | M2                       | X2                            | 5.4                     | 49.9                      |
| 11  | 12  | 1     | 0.333                      | 0.333                          | 0.333                     | M1                       | X3                            | 7.2                     | 50.2                      |
| 12  | 9   | 1     | 0.000                      | 0.500                          | 0.500                     | M2                       | X1                            | 9.0                     | 48.2                      |
| 13  | 14  | 1     | 0.000                      | 1.000                          | 0.000                     | M1                       | X2                            | 3.6                     | 46.7                      |
| 14  | 33  | 1     | 0.500                      | 0.500                          | 0.000                     | M2                       | X3                            | 12.0                    | 48.2                      |
| 15  | 2   | 1     | 0.500                      | 0.000                          | 0.500                     | M1                       | X3                            | 6.2                     | 49.9                      |
| 16  | 20  | 1     | 0.000                      | 0.500                          | 0.500                     | M1                       | X3                            | 5.3                     | 47.9                      |
| 17  | 34  | 1     | 0.500                      | 0.000                          | 0.500                     | M2                       | X1                            | 10.0                    | 50.2                      |
| 18  | 21  | 1     | 0.500                      | 0.000                          | 0.500                     | M2                       | X2                            | 5.3                     | 49.2                      |
| 19  | 15  | 1     | 0.000                      | 0.500                          | 0.500                     | M2                       | X2                            | 2.4                     | 50.0                      |
| 20  | 7   | 1     | 0.000                      | 1.000                          | 0.000                     | M2                       | X3                            | 6.6                     | 47.8                      |

(continued)

TABLE E13.6 (Continued)

| Std | Run | Block | Component 1 |             | Component 2<br>Crosslinker (%) | Component 3,<br>Resin (%) | Factor 1,   |                  | Factor 2, |          | Response 1, |            | Response 2, |  |
|-----|-----|-------|-------------|-------------|--------------------------------|---------------------------|-------------|------------------|-----------|----------|-------------|------------|-------------|--|
|     |     |       | Monomer (%) | Monomer (%) |                                |                           | Vendor Type | Crosslinker Type | Hardness  | Hardness | Solids (%)  | Solids (%) |             |  |
| 21  | 5   | 1     | 0.333       | 0.333       | 0.333                          | 0.333                     | M2          | X3               | X3        | 9.0      | 9.0         | 50.3       | 50.3        |  |
| 22  | 31  | 1     | 0.500       | 0.500       | 0.500                          | 0.000                     | M1          | X2               | X2        | 7.0      | 7.0         | 46.9       | 46.9        |  |
| 23  | 29  | 1     | 0.000       | 0.500       | 0.500                          | 0.500                     | M1          | X2               | X2        | 3.5      | 3.5         | 47.8       | 47.8        |  |
| 24  | 3   | 1     | 1.000       | 0.000       | 0.000                          | 0.000                     | M1          | X3               | X3        | 5.0      | 5.0         | 48.7       | 48.7        |  |
| 25  | 28  | 1     | 0.500       | 0.000       | 0.000                          | 0.500                     | M1          | X2               | X2        | 5.0      | 5.0         | 49.9       | 49.9        |  |
| 26  | 22  | 1     | 0.000       | 0.000       | 0.000                          | 1.000                     | M2          | X1               | X1        | 7.7      | 7.7         | 50.9       | 50.9        |  |
| 27  | 16  | 1     | 0.500       | 0.500       | 0.500                          | 0.000                     | M1          | X3               | X3        | 8.7      | 8.7         | 48.0       | 48.0        |  |
| 28  | 13  | 1     | 1.000       | 0.000       | 0.000                          | 0.000                     | M2          | X2               | X2        | 6.1      | 6.1         | 49.0       | 49.0        |  |
| 29  | 8   | 1     | 0.000       | 1.000       | 1.000                          | 0.000                     | M2          | X2               | X2        | 5.9      | 5.9         | 47.2       | 47.2        |  |
| 30  | 1   | 1     | 1.000       | 0.000       | 0.000                          | 0.000                     | M1          | X1               | X1        | 11.0     | 11.0        | 47.8       | 47.8        |  |
| 31  | 30  | 1     | 1.000       | 0.000       | 0.000                          | 0.000                     | M2          | X3               | X3        | 10.0     | 10.0        | 47.8       | 47.8        |  |
| 32  | 38  | 1     | 0.000       | 1.000       | 1.000                          | 0.000                     | M1          | X3               | X3        | 6.3      | 6.3         | 46.8       | 46.8        |  |
| 33  | 36  | 1     | 0.500       | 0.500       | 0.500                          | 0.000                     | M2          | X1               | X1        | 14.0     | 14.0        | 46.9       | 46.9        |  |
| 34  | 25  | 1     | 0.000       | 0.500       | 0.500                          | 0.500                     | M2          | X1               | X1        | 9.2      | 9.2         | 50.4       | 50.4        |  |
| 35  | 37  | 1     | 0.000       | 0.000       | 0.000                          | 1.000                     | M2          | X3               | X3        | 7.7      | 7.7         | 49.4       | 49.4        |  |
| 36  | 11  | 1     | 0.333       | 0.333       | 0.333                          | 0.333                     | M2          | X2               | X2        | 6.7      | 6.7         | 50.3       | 50.3        |  |
| 37  | 10  | 1     | 0.000       | 0.000       | 0.000                          | 1.000                     | M1          | X1               | X1        | 7.3      | 7.3         | 50.9       | 50.9        |  |
| 38  | 6   | 1     | 0.000       | 0.000       | 0.000                          | 1.000                     | M2          | X2               | X2        | 5.0      | 5.0         | 49.9       | 49.9        |  |

- 13.21** Consider a mixture experiment with the following constraints on the component proportions:

$$\begin{aligned}x_1 &\leq 0.2, & x_2 &\leq 0.6, & x_3 &\leq 0.15, \\x_4 &\leq 0.65, & x_5 &\leq 0.15, & x_6 &\leq 0.75 \\x_1 + x_2 + x_3 + x_4 + x_5 + x_6 &= 1\end{aligned}$$

- (a) Suggest a screening design for this mixture experiment.  
(b) Suppose that you have obtained the following first-order model for the response.

$$\hat{y} = 20x_1 + 30x_2 + 10x_3 + 16x_4 + 45x_5 + 80x_6$$

Calculate the adjusted effects of each component. Interpret the results of this analysis.

- 13.22** Consider a mixture experiment with the following constraints on the component proportions.

$$0.05 \leq x_1, \quad 0.05 \leq x_2, \quad 0.10 \leq x_3,$$

$$0.20 \leq x_4, \quad 0.01 \leq x_5$$

$$x_1 + x_2 + x_3 + x_4 + x_5 = 1$$

- (a) Suggest a screening design for this mixture experiment.  
(b) Suppose that you have obtained the following first-order model for the response:

$$\hat{y} = 20x_1 + 65x_2 + 20x_3 + 90x_4 + 45x_5$$

Calculate the adjusted effects of each component. Interpret the results of this analysis.

- 13.23** Consider the mixture-process experiment of Example 13.6. We now assume both mixing time and plodder temperature are considered noise variables.

- (a) Find the estimated mean response surface for throughput as a function of the mixture variables.  
(b) Find the estimated variance response surface for throughput as a function of  $\sigma_{z_1}^2$  and  $\sigma_{w_1}^2$ .  
(c) Find the combination of mixture components to simultaneously maximize average throughput and minimize variability.

- 13.24** Show why the main effects for the process variables can not be included in the model given in Equation 13.10.

- 13.25** Construct a contour plot of the scaled prediction variance in the design region for the  $D$ -optimal design in Table 13.8 used for the shampoo foam experiment. Where are the worst regions for prediction? Where are the best?

- 13.26** Consider the three-component mixture experiment in Exercise 13.2. Find a 10-run *I*-optimal design for this problem. Compare this design to the one used in Exercise 13.2 in terms of prediction variance.
- 13.27** Consider the three-component mixture experiment in Exercise 13.3. Find a 10-run *I*-optimal design for this problem. Compare this design to the one used in Exercise 13.3 in terms of prediction variance.
- 13.28** Consider the mixture problem described in Exercise 13.4, but suppose that one of the constraints is modified as follows:  $0.15 \leq x_1 \leq 0.50$ . You plan to fit a quadratic mixture model.
- (a) Graph the mixture space.
  - (b) Suppose that you will consider using 15 runs in this experiment. Construct an *I*-optimal design.
  - (c) Construct a 15-run *D*-optimal design.
  - (d) Construct a 15-run space-filling design.
  - (e) Compare these designs in terms of prediction variance. Which design do you prefer?
- 13.29** Reconsider the situation in Exercise 13.28. Suppose after running the experiment you find out that it is necessary to fit a special cubic model to the response. Evaluate the three designs that you constructed in Exercise 13.28 with respect to fitting the special cubic model. Is any one of the three designs preferable? Does this tell you anything about how robust the choice of design is to the form of the model that is actually fit to the data?
- 13.30** Reconsider the four-component mixture experiment described in Exercise 13.14.
- (a) Construct a 20-run *I*-optimal design for this problem.
  - (b) Construct a 20-run *D*-optimal design for this problem.
  - (c) Construct a 20-run space-filling design for this problem.
  - (d) Compare the design you have constructed in parts (a)–(c) with the original design in Exercise 13.14. Which design would you prefer?

### Short Answer Questions

- 13.31** Using ratios of components to design a mixture experiment allows the experimenter to use a standard mixture design that will fill the original mixture space. **True False**
- 13.32** Upper bound constraints in a mixture experiment always leave the original simplex design region unchanged. **True False**
- 13.33** Using ratios of components is an approach that can be used for including process variables in mixture experiments. **True False**
- 13.34** Rotatability is a criterion that is useful for designing mixture experiments. **True False**
- 13.35** Response trace plots can be useful in screening mixture components. **True False**

---

## APPENDIX 1

---

### MOMENT MATRIX OF A ROTATABLE DESIGN

In the development that follows, it is convenient in expressing the polynomial model making use of **derived power vectors** and **Schlaifflian matrices**. If  $\mathbf{z}' = [z_1, z_2, \dots, z_k]$ , then  $\mathbf{z}'^{[p]}$ , the derived power vector of degree  $p$ , is defined such that

$$\mathbf{z}'^{[p]} \mathbf{z}^{[p]} = (\mathbf{z}' \mathbf{z})^p$$

For example if  $\mathbf{z}' = [z_1, z_2, z_3]$ , then

$$\mathbf{z}'^{[2]} = \left[ z_1^2, z_2^2, z_3^2, \sqrt{2}z_1z_2, \sqrt{2}z_1z_3, \sqrt{2}z_2z_3 \right]$$

and

$$\mathbf{z}'^{[1]} = \mathbf{z}'$$

If a vector  $\mathbf{x}$  is formed from a vector  $\mathbf{z}$  containing  $k$  elements through the transformation

$$\mathbf{x} = \mathbf{H}\mathbf{z} \tag{A.1.1}$$

then the **Schlaifflian matrix**  $\mathbf{H}^{[p]}$  is defined such that

$$\mathbf{x}^{[p]} = \mathbf{H}^{[p]} \mathbf{z}^{[p]}$$

---

*Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, Fourth Edition.  
by Raymond H. Myers, Douglas C. Montgomery, and Christine M. Anderson-Cook.  
Copyright © 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

It is readily seen that if the transformation matrix  $\mathbf{H}$  is orthogonal, then  $\mathbf{H}^{[p]}$  is also orthogonal. One can write

$$\mathbf{x}'^{[p]}\mathbf{x}^{[p]} = \mathbf{z}'^{[p]}\mathbf{H}'^{[p]}\mathbf{H}^{[p]}\mathbf{z}^{[p]} \quad (\text{A.1.2})$$

the left-hand side of Equation A.1.2 is, by definition,  $(\mathbf{x}'\mathbf{x})^p$ . Because  $\mathbf{H}$  is orthogonal,

$$(\mathbf{x}'\mathbf{x})^p = (\mathbf{z}'\mathbf{z})^p = \mathbf{z}'^{[p]}\mathbf{z}^{[p]}$$

and thus the Schlaiffian matrix  $\mathbf{H}^{[p]}$  is orthogonal. Another result that is quite useful in what follows is that, given two vectors  $\mathbf{x}$  and  $\mathbf{z}$ , each having  $k$  elements, then

$$(\mathbf{x}'\mathbf{z})^p = \mathbf{x}'^{[p]}\mathbf{z}^{[p]}$$

For a response function of order  $d$ , the estimated response  $\hat{y}$  can be written in the form

$$\hat{y} = \mathbf{x}'^{[d]}\mathbf{b} \quad (\text{A.1.3})$$

where for a point  $(x_1, x_2, \dots, x_k)$  we have

$$\mathbf{x}' = (1, x_1, x_2, \dots, x_k)$$

and the vector  $\mathbf{b}$  contains the least squares estimators  $b_0, b_1, \dots$ , for the vector  $\mathbf{x}^{[d]}$  with suitable multipliers. For example, for  $k=2$  and  $d=2$ ,  $\mathbf{x}' = [1, x_1, x_2]$ , and  $\mathbf{b}'$  and  $\mathbf{x}'^{[2]}$  are given by

$$\begin{aligned} \mathbf{b}' &= [b_0, b_1/\sqrt{2}, b_2/\sqrt{2}, b_{11}, b_{22}, b_{12}/\sqrt{2}] \\ \mathbf{x}'^{[2]} &= [1, \sqrt{2}x_1, \sqrt{2}x_2, x_1^2, x_2^2, \sqrt{2}x_1x_2] \end{aligned}$$

Thus, from Equation A.1.3 we obtain

$$\begin{aligned} \text{Var}[\hat{y}(\mathbf{x})] &= \mathbf{x}'^{[d]}[\text{Var}(\mathbf{b})]\mathbf{x}^{[d]} \\ &= \sigma^2 \mathbf{x}'^{[d]}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}^{[d]} \end{aligned} \quad (\text{A.1.4})$$

where  $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$  is the variance-covariance matrix of the vector  $\mathbf{b}$ .

Consider now a *second* point  $(z_1, z_2, \dots, z_k)$ , which is the same distance from the origin as the point described by  $(x_1, x_2, \dots, x_k)$ . Denote by  $\mathbf{z}'$  the vector  $(1, z_1, z_2, \dots, z_k)$ . There is, then, an orthogonal matrix  $\mathbf{R}$  for which

$$\mathbf{z} = \mathbf{R}\mathbf{x} \quad (\text{A.1.5})$$



where  $\mathbf{R}$  is of the form

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & & & & \\ 0 & & & & \\ \vdots & & & \mathbf{H}_{k \times k} & \\ 0 & & & & \end{bmatrix} \tag{A.1.6}$$

and  $\mathbf{H}$  is an orthogonal matrix with the dimensions indicated in Equation A.1.6. The variance of the estimated response at the second point is then

$$\text{Var}[\hat{y}(\mathbf{z})] = \sigma^2 \mathbf{z}'^{[d]} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{z}^{[d]}$$

Let  $\mathbf{R}^{[d]}$  be the Schlaifflian matrix of the transformation in Equation A.1.5:

$$\begin{aligned} \text{Var}[\hat{y}(\mathbf{z})] &= \sigma^2 \mathbf{x}'^{[d]} \mathbf{R}'^{[d]} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{R}^{[d]} \mathbf{x}^{[d]} \\ &= \sigma^2 \mathbf{x}'^{[d]} [\mathbf{R}'^{[d]} (\mathbf{X}'\mathbf{X}) \mathbf{R}^{[d]}]^{-1} \mathbf{x}^{[d]} \end{aligned} \tag{A.1.7}$$

because  $\mathbf{R}^{[d]}$  is orthogonal. For the design to be rotatable,  $\text{Var}(\hat{y})$  is constant on spheres, which implies that for any orthogonal matrix  $\mathbf{H}$  we have

$$\mathbf{X}'\mathbf{X} = \mathbf{R}'^{[d]} (\mathbf{X}'\mathbf{X}) \mathbf{R}^{[d]} \tag{A.1.8}$$

where  $\mathbf{R}$  is of the form indicated in Equation A.1.6. The requirement in Equation A.1.8 essentially means that the moment matrix remains the same if the design is **rotated**—that is, if the rows of the *design* matrix, denoted by  $\mathbf{D}$  in the equation

$$\begin{bmatrix} 1 & & & & \\ 1 & & & & \\ \vdots & & & & \\ 1 & & & & \end{bmatrix} \mathbf{D} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{1N} & x_{2N} & \cdots & x_{kN} \end{bmatrix} = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_N \end{bmatrix} \tag{A.1.9}$$

are rotated via the transformation

$$\mathbf{z}_i = \mathbf{R}' \mathbf{x}_i$$

It is easily seen that the rotated design will have moment matrix (apart from the constant  $N^{-1}$ ) equal to the right-hand side of Equation A.1.8.

Consider now a vector  $\mathbf{t}' = [1, t_2, t_1, \dots, t_k]$  of dummy variables. The utility of these variables is in the construction of a generating function for the design moments. Consider the quantity

$$\text{MGF} = N^{-1} \mathbf{t}'^{[d]} \mathbf{X}' \mathbf{X} \mathbf{t}^{[d]} \tag{A.1.10}$$

The matrix  $\mathbf{X}'\mathbf{X}$  is alternatively given by

$$\mathbf{X}'\mathbf{X} = \sum_{u=1}^N \mathbf{x}_u^{[d]} \mathbf{x}_u'^{[d]}$$

where the vector  $\mathbf{x}_u' = [1, x_{1u}, x_{2u}, \dots, x_{ku}]$  refers to the  $u$ th row of the design matrix, augmented by 1—that is, the  $u$ th row of the matrix in Equation A.1.9. Thus

$$\begin{aligned} \text{MGF} &= N^{-1} \sum_{u=1}^N \{ \mathbf{t}'^{[d]} \mathbf{x}_u^{[d]} \mathbf{x}_u'^{[d]} \mathbf{t}^{[d]} \} \\ &= N^{-1} \sum_{u=1}^N \{ \mathbf{t}' \mathbf{x}_u \}^{2d} \end{aligned} \quad (\text{A.1.11})$$

From Equation A.1.11, it is seen that upon expanding  $\mathbf{t}' \mathbf{x}_u$  we have

$$\text{MGF} = N^{-1} \sum_{u=1}^N \{ 1 + t_1 x_{1u} + t_2 x_{2u} + \dots + t_k x_{ku} \}^{2d} \quad (\text{A.1.12})$$

When Equation A.1.12 is expanded, the terms involve moments of the design through order  $2d$ . In fact, the coefficient of  $t_1^{\delta_1} t_2^{\delta_2} \dots t_k^{\delta_k}$  is

$$\frac{(2d)!}{\prod_{i=1}^k (\delta_i)! (2d - \delta)!} [1^{\delta_1} 2^{\delta_2} \dots k^{\delta_k}] \quad (\text{A.1.13})$$

where  $\sum_{i=1}^k \delta_i = \delta \leq 2d$ . For a rotatable design,

$$\begin{aligned} \text{MGF} &= N^{-1} \mathbf{t}'^{[d]} (\mathbf{X}'\mathbf{X}) \mathbf{t}^{[d]} = N^{-1} \mathbf{t}'^{[d]} \mathbf{R}'^{[d]} \mathbf{X}' \mathbf{X} \mathbf{R}^{[d]} \mathbf{t}^{[d]} \\ &= N^{-1} (\mathbf{t}' \mathbf{R}')^{[d]} \mathbf{X}' \mathbf{X} [\mathbf{R} \mathbf{t}]^{[d]} \end{aligned}$$

where  $\mathbf{R}$  is the  $(k+1) \times (k+1)$  orthogonal matrix introduced in Equation A.1.6. This implies that for a rotatable design, an orthogonal transformation on  $\mathbf{t}$  does not affect the MGF. Because the MGF is a polynomial in the  $t$ 's (also involving the design moments), for a rotatable design, the MGF must be a function of  $\sum_{i=1}^k t_i^2$ . That is, it is of the form

$$\text{MGF} = \sum_{j=0}^d a_{2j} \left[ \sum_{i=1}^k t_i^2 \right]^j \quad (\text{A.1.14})$$

It is easily seen that the coefficient of  $t_1^{\delta_1} t_2^{\delta_2} \dots t_k^{\delta_k}$  in Equation A.1.14 is zero *if any of the  $\delta_i$  are odd*. For the case where all  $\delta_i$  are even, the coefficient from the multinomial expansion of  $\left[ \sum_{i=1}^k t_i^2 \right]^j$  is given by

$$\frac{a_\delta (\delta/2)!}{\prod_{i=1}^k (\delta_i/2)!} \tag{A.1.15}$$

The reader should now consider Equation A.1.15 in conjunction with Equation A.1.13, the former pertaining to the generating function for the moments *in general*, and the latter pertaining to the case of a rotatable design, with the value being zero for moments with any  $\delta_i$  odd. Upon equating the two and solving for the moment, the result is

$$N^{-1} \sum_{u=1}^N x_{1u}^{\delta_1} x_{2u}^{\delta_2} \dots x_{ku}^{\delta_k} = \frac{\lambda_\delta \prod_{i=1}^k (\delta_i)!}{2^{\delta/2} \prod_{i=1}^k (\delta_i/2)!} \tag{A.1.16}$$

for all  $\delta_i$  even, and

$$N^{-1} \sum_{u=1}^N x_{1u}^{\delta_1} x_{2u}^{\delta_2} \dots x_{ku}^{\delta_k} = 0 \tag{A.1.17}$$

for any  $\delta_i$  odd. Here,  $\lambda_\delta$  is given by

$$\lambda_\delta = \frac{a_\delta 2^{\delta/2} (\delta/2)! (2d - \delta)!}{(2d)!} \tag{A.1.18}$$

If we consider Equations A.1.16 and A.1.18 for the second-order case, we have  $d=2$ , and thus  $[i] = [ij] = [ijk] = [iij] = [iii] = 0$  for  $i \neq j \neq k$ ,  $[ii] = \lambda_2$  (fixed by scaling) and  $[iii]/[iij] = 3$ .



---

## APPENDIX 2

---

### ROTABILITY OF A SECOND-ORDER EQUIRADIAL DESIGN

Consider first the moment of  $x_1$  of order  $\delta$  ( $\delta = 1, 2, 3, 4$ ). After multiplication by  $n_1$ , we obtain

$$\sum_{u=0}^{n_1-1} x_{1u}^\delta = \sum_{u=0}^{n_1-1} \left\{ \rho \cos \left( \theta + \frac{2\pi u}{n_1} \right) \right\}^\delta \quad (\text{A.2.1})$$

Because  $\cos \tau = [e^{i\tau} + e^{-i\tau}]/2$ , Equation A.2.1 can be written

$$\sum_{u=0}^{n_1-1} x_{1u}^\delta = \left( \frac{\rho}{2} \right)^\delta \sum_{u=0}^{n_1-1} [a\omega^u + a^{-1}\omega^{-u}]^\delta$$

where  $\omega = e^{2i\pi/n_2}$  and  $a = e^{i\theta}$ . Using a binomial expansion,

$$\sum_{u=0}^{n_1-1} x_{1u}^\delta = \left( \frac{\rho}{2} \right)^\delta \sum_{t=0}^{\delta} \binom{\delta}{t} a^{\delta-2t} \sum_{u=0}^{n_1-1} \omega^{(\delta-2t)u} \quad (\text{A.2.2})$$

An expression similar to Equation A.2.2 for the moments of  $x_2$  can be established with little difficulty. In considering the portion  $\sum_{u=0}^{n_1-1} \omega^{(\delta-2t)u}$  of Equation A.2.2 for  $t = 0, 1, 2, \dots, \delta$ , it is first noted that

$$-\delta \leq \delta - 2t \leq \delta$$

The interest here is in moments of order  $n_1 - 1$  and less, because  $n_1 \geq 5$ . If  $\delta - 2t = 0$ , then

$$\sum_{u=0}^{n_1-1} \omega^{(\delta-2t)u} = n_1$$

On the other hand, it is not difficult to show that for  $|\delta - 2t| \leq n_1 - 1$  and nonzero, we obtain

$$\sum_{u=0}^{n_1-1} \omega^{(\delta-2t)u} = 0$$

Using Equation A.2.2 for  $\delta = 1$ ,  $\delta - 2t$  takes on values  $-1$  and  $+1$  and thus  $\sum_{u=0}^{n_1-1} x_{1u} = 0$ . Likewise, for the case of  $\delta = 3$ , it is found that  $\sum_{u=0}^{n_1-1} x_{1u}^3 = 0$ . In a similar fashion, it can be shown that the odd moments of  $x_2$  are zero. For  $\delta = 2$ ,  $\delta - 2t$  does that on a zero value when  $t = 1$ . Evaluating Equation A.2.2

$$n_1[11] = \frac{n_1 \binom{2}{1} \rho^2}{4}$$

and thus

$$[11] = \frac{\rho^2}{2}$$

For  $\delta = 4$ , the nonzero contribution in Equation A.2.2 appears when  $t = 2$ . Thus the pure fourth moment is given by

$$\begin{aligned} [1111] &= \frac{\binom{4}{2} \rho^4}{16} \\ &= \frac{3\rho^4}{8} \end{aligned}$$

Similar procedures can be used for the case of  $x_2$  to show that

$$n_1[2^2] = \frac{\rho^2 n_1}{2}, \quad n_1[2^4] = \frac{\rho^4 n_1}{8}$$

At this point, consider moments of the type  $\sum_{u=0}^{n_1-1} x_{1u}^{\delta_1} x_{2u}^{\delta_2}$ , where  $\delta_1 + \delta_2 \leq 4$ . One can write

$$\begin{aligned} \sum_{u=0}^{n_1-1} x_{1u}^{\delta_1} x_{2u}^{\delta_2} &= \sum_{u=0}^{n_1-1} \left\{ \rho \cos \left( \theta + \frac{2\pi u}{n_1} \right) \right\}^{\delta_1} \left\{ \rho \sin \left( \theta + \frac{2\pi u}{n_1} \right) \right\}^{\delta_2} \\ &= \left( \frac{\rho}{2} \right)^{\delta_1 + \delta_2} \left( \frac{1}{i} \right)^{\delta_2} \sum_{u=0}^{n_1-1} \{ a\omega^u + a^{-1}\omega^{-u} \}^{\delta_1} \{ a\omega^u - a^{-1}\omega^{-u} \}^{\delta_2} \end{aligned} \tag{A.2.3}$$

where  $a$  and  $\omega$  are as before and  $i = \sqrt{-1}$ . For  $\delta_1 = 1$  and  $\delta_2 = 1$ , Equation A.2.3 is easily seen to be zero after making use of the fact that  $\sum_{u=0}^{n_1-1} \omega^{2u}$  and  $\sum_{u=0}^{n_1-1} \omega^{-2u} = 0$ . Likewise,  $n_1[112]$  and  $n_1[122]$  are found to be zero. For  $n_1[1112]$ , Equation A.2.3 becomes

$$\begin{aligned} & \left(\frac{\rho}{2}\right)^4 \left(\frac{1}{i}\right) \sum_{u=0}^{n_1-1} \{a^3 \omega^{3u} + a^{-3} \omega^{-3u} + 3a\omega^u + 3a^{-1} \omega^{-u}\} \{a\omega^u - a^{-1} \omega^{-u}\} \\ &= \left(\frac{\rho}{2}\right)^4 \left(\frac{1}{i}\right) \{-3n_1 + 3n_1\} \\ &= 0 \end{aligned}$$

Similarly  $n_1[1222]$  can be shown to be zero.

It remains now to develop the expression for  $\sum_{u=0}^{n_1-1} x_{1u}^2 x_{2u}^2$ . We can once again use Equation A.2.3 to develop the following:

$$\begin{aligned} \sum_{u=0}^{n_1-1} x_{1u}^2 x_{2u}^2 &= \left(\frac{\rho}{2}\right)^4 (-1) \sum_{u=0}^{n_1-1} \{a^2 \omega^{2u} + 2 + a^{-2} \omega^{-2u}\} \{a^2 \omega^{2u} - 2 + a^{-2} \omega^{-2u}\} \\ &= \frac{\rho^4}{16} (-1) \{n_1 - 4n_1 + n_1\} \\ &= \frac{\rho^4 n_1}{8} \end{aligned}$$

On the basis of  $[iiii]$  and  $[ijij]$  in the above developments, it is clear that the equiradial design is rotatable.





## REFERENCES

- Admassu, W., and Breese, T. (1999), "Feasibility of Using Natural Fishbone Apatite as a Substitute for Hydroxyapatite in Remediating Aqueous Heavy Metals," *Journal of Hazardous Materials*, 69, 187–196.
- Ahmadi, M., Vahabzadeh, F., Bonakdarpour, B., Mofarrah, E., and Mehranian, M. (2005), "Application of the Central Composite Design and Response Surface Methodology to the Advanced Treatment of Olive Oil Processing Wastewater Using Fenton's Peroxidation," *Journal of Hazardous Materials*, B123, 187–195.
- Aitken, M. (1987), "Modeling Variance Heterogeneity in Normal Regression Using GLIM," *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 36, 332–339.
- Allen, D. M. (1971), "Mean Square Error of Prediction as a Criterion for Selecting Variables," *Technometrics*, 13, 469–475.
- Allen, D. M. (1974), "The Relationship Between Variable Selection and Data Augmentation and a Method for Prediction," *Technometrics*, 16, 125–127.
- Allen, T. T., Yu, L., and Schmitz, J. (2003), "An Experimental Design Criterion for Minimizing Meta-Model Prediction Errors Applied to a Die Casting Process Design," *Applied Statistics*, 52, 103–117.
- Amenaghawon, N. A., Nwaru, K. I., Aisien, F. A., Ogbeide, S. E., and Okieimen, C. O. (2013), "Application of Box-Behnken Design for the Optimization of Citric Acid Production from Corn Starch Using *Aspergillus niger*," *British Biotechnology Journal*, 3, 236–245.
- Ames, A. E., Mattucci, D., MacDonald, S., Szongi, G., and Hawkins, D. M. (1997), "Quality Loss Functions for Optimization Across Multiple Response Surfaces," *Journal of Quality Technology*, 29, 339–346.
- Anderson-Cook, C. M., Goldfarb, H., Borror, C. M., Montgomery, D. C., Canter, K. G., and Twist, J. A., (2004), "Mixture and Mixture Process Variables Experiments for Pharmaceutical Applications," *Pharmaceutical Statistics*, 3, 247–260.

- Anderson-Cook, C. M., Borror, C. M., and Jones, B. (2009), "Graphical Tools for Assessing the Sensitivity of Response Surface Designs to Model Misspecification," *Technometrics*, 51, 75–87.
- Anderson-Cook, C. M., Borror, C. M., and Montgomery, D. C. (2009a), "Response Surface Design Evaluation and Comparison" (with discussion), *Journal of Statistical Planning and Inference*, 139, 629–674.
- Anderson-Cook, C. M., Borror, C. M., and Jones, B. (2009b), "Graphical Tools for Assessing the Sensitivity of Response Surface Designs to Model Misspecification," *Technometrics*, 51, 75–87.
- Anderson-Cook, C. M., and Lu, L. (2015), "Much-Needed Structure: A New 5-Step Decision-Making Process Helps You Evaluate, Balance Competing Objectives," *Quality Progress*, 48(10), 42–50.
- Barghash, M. A. and Alkaabneh, F. A. (2014), "Shrinkage and Warpage Detailed Analysis and Optimization for the Injection Molding Process Using Multistage Experimental Design," *Quality Engineering*, 26, 319–334.
- Bartlett, M. S., and Kendall, D. G. (1946), "The Statistical Analysis of Variance Heterogeneity and the Logarithmic Transformation," *Journal of the Royal Statistical Society, Series B*, 8, 128–150.
- Belsley, D. A., Kah, E., and Welsch, R. E. (1980), *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, John Wiley & Sons, New York.
- Bettonvil, B., and Kleijnen, J. J. C. (1996), "Searching for Importance Factors in Simulation Models with Many Factors: Sequential Bifurcation," *European Journal of Operational Research*, 96, 180–194.
- Biles, W. E. (1975), "A Response Surface Method for Experimental Optimization of Multiresponse Processes," *Industrial and Engineering Chemistry—Process Design and Development*, 14, 152–158.
- Bisgaard, S. (2000), "The Design and Analysis of  $2^{k-p} \times 2^{q-r}$  Split-Plot Experiments," *Journal of Quality Technology*, 32, 39–56.
- Bisgaard, S., and Ankenman, B. (1996), "Standard Errors for the Eigenvalues in Second-order Response Surface Models," *Technometrics*, 38, 238–246.
- Bishop, C. M. (1995), *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford, UK.
- Booth, K. H. V., and Cox, D. R. (1962), "Some Systematic Supersaturated Designs," *Technometrics*, 4, 489–495.
- Borkowski, J. J., and Lucas, J. M. (1997), "Designs of Mixed Resolution for Process Robustness Studies," *Technometrics* 39, 63–70.
- Borror, C. M. (1998), "Response Surface Methods for Experiments Involving Noise Variables," Ph.D. Dissertation, Department of Industrial Engineering, Arizona State University, Tempe, AZ.
- Borror, C. M., and Montgomery, D. C. (2000), "Mixed Resolution Designs as Alternatives to Taguchi Inner/Outer Array Designs for Robust Design Problems," *Quality and Reliability Engineering International*, 16, 117–127.
- Borror, C. M., Montgomery, D. C., and Myers, R. H. (2002), "Evaluation of Statistical Designs for Experiments with Noise Variables," to appear in the *Journal of Quality Technology*.
- Box, G. E. P. (1963), "The Effect of Errors in the Factor Levels and Experimental Design," *Technometrics*, 6, 247–262.
- Box, G. E. P. (1988), "Signal-to-Noise Ratios, Performance Criteria, and Transformations" (with discussion), *Technometrics*, 30, 1–40.
- Box, G. E. P., and Behnken, D. W. (1960), "Some New Three-Level Designs for the Study of Quantitative Variables," *Technometrics*, 2, 455–475.
- Box, G. E. P., and Cox, D. R. (1964), "An Analysis of Transformations" (with discussion), *Journal of the Royal Statistical Society, Series B*, 26, 211–246.
- Box, G. E. P., and Draper, N. R. (1959), "A Basis for the Selection of a Response Surface Design," *Journal of the American Statistical Association*, 54, 622–654.

- Box, G. E. P., and Draper, N. R. (1963), "The Choice of a Second Order Rotatable Design," *Biometrika*, 50, 335–352.
- Box, G. E. P., and Draper, N. R. (1971), "Factorial Designs, the X'X Criterion, and Some Related Matters," *Technometrics*, 13, 731–742.
- Box, G. E. P., and Draper, N. R. (1974), "Some Minimum Point Designs for Second Order Response Surfaces," *Technometrics*, 16, 613–616.
- Box, G. E. P., and Draper, N. R. (1975), "Robust Designs," *Biometrika*, 62, 347–352.
- Box, G. E. P., and Draper, N. R. (1987), *Empirical Model-Building and Response Surfaces*, John Wiley & Sons, New York.
- Box, G. E. P., and Draper, D. R. (2007), *Response Surfaces, Mixtures and Ridge Analyses*, Wiley and Sons, New York.
- Box, G. E. P., and Hunter, J. S. (1954), "A Confidence Region for the Solution of a Set of Simultaneous Equations with an Application to Experimental Design," *Biometrika*, 41, 190–199.
- Box, G. E. P., and Hunter, J. S. (1957), "Multifactor Experimental Designs for Exploring Response Surfaces," *The Annals of Mathematical Statistics*, 28, 195–241.
- Box, G. E. P., and Hunter, J. S. (1961a), "The  $2^{k-p}$  Fractional Factorial Designs, Part I," *Technometrics*, 2, 311–352.
- Box, G. E. P., and Hunter, J. S. (1961b), "The  $2^{k-p}$  Fractional Factorial Designs, Part II," *Technometrics*, 3, 449–458.
- Box, G. E. P., and Jones, S. (1989), "Designing Products that are Robust to the Environment," paper presented at ASA Conference, Washington, DC.
- Box, G. E. P., and Jones, S. (1992), "Split-Plot Designs for Robust Product Experimentation," *Journal of Applied Statistics*, 19, 3–26.
- Box, G. E. P. and Jones, S. (2000), "Split Plots for Robust Product and Process Experimentation," *Quality Engineering*, 13, 127–134.
- Box, G. E. P., and Meyer, R. D. (1986), "Dispersion Effects from Fractional Designs," *Technometrics*, 28, 19–27.
- Box, G. E. P., and Wetz, J. M. (1973), "Criterion for Judging the Adequacy of Estimation by an Approximation Response Polynomial," Technical Report No. 9, Department of Statistics, University of Wisconsin, Madison.
- Box, G. E. P., and Wilson, K. B. (1951), "On the Experimental Attainment of Optimum Conditions," *Journal of the Royal Statistical Society, Series B*, 13, 1–45.
- Brent, R. P. (1973), "Some Efficient Algorithms for Solving Systems of Nonlinear Equations," *SIAM Journal on Numerical Analysis*, 10, 327–344.
- Bursztyn, D., and Steinberg, D. M. (2006), "Comparison of Designs for Computer Experiments," *Journal of Statistical Planning and Inference*, 136, 1103–1119.
- Byrne, D. M., and Taguchi, G. (1987), "The Taguchi Approach to Parameter Design," *Quality Progress*, December, 1987, 19–26.
- Carlyle, W. M., Montgomery, D. C., and Runger, G. C. (2000), "Optimization Problems and Methods in Quality Control and Improvement," (with discussion), *Journal of Quality Technology*, 31, 1–17.
- Carpenter, W. C., and Barthelemy, J. (1993), "A Comparison of Polynomial Approximations and Artificial Neural Nets as Response Surfaces," *Structural Optimization*, 5, 166–174.
- Carroll, R. J., and Ruppert, D. (1988), *Transformation and Weighting in Regression*, Chapman and Hall, New York.
- Carter, W. H., Jr., Chinchilli, V. M., and Campbell, E. D. (1990), "A Large-Sample Confidence Region Useful in Characterizing the Stationary Point of a Quadratic Response Surface," *Technometrics*, 32, 425–435.

- Chapman, J., Lu, L., and Anderson-Cook, C. M. (2014a), "Process Optimization for Multiple Responses Utilizing the Pareto Front Approach," *Quality Engineering*, 26, 253–268.
- Chapman, J. L., Lu, L., and Anderson-Cook, C. M. (2014b), "Incorporating Response Variability and Estimation Uncertainty into Pareto Front Optimization," *Computers & Industrial Engineering*, 76, 253–267.
- Chapman, J. L., Lu, L., and Anderson-Cook, C. M. (2015), "Impact of Response Variability on Pareto Front Optimization," *Statistical Analysis and Data Mining*, 8, 314–328.
- Cheng, B., and Titterton, D. M. (1994), "Neural Networks: A Review from a Statistical Perspective," *Statistical Science*, 9, 2–54.
- Chung, P. J., Goldfarb, H. B., and Montgomery, D. C. (2007), "Optimal Designs for Mixture-Process Experiments with Control and Noise Variables," *Journal of Quality Technology*, 39, 179–190.
- Condra, L. W. (1993), *Reliability Improvement with Design of Experiments*, Marcel Dekker, New York.
- Cook, R. D. (1977), "Detection of Influential Observation in Linear Regression," *Technometrics*, 19, 15–17.
- Cook, R. D. (1979), "Influential Observations in Linear Regression," *Journal of the American Statistical Association*, 74, 169–174.
- Cornell, J. A. (1986), "A Comparison Between Two Ten-Point Designs for Studying Three-Component Mixture Systems," *Journal of Quality Technology*, 18, 1–15.
- Cornell, J. A. (1995), "Fitting Models to Data from Mixture Experiments Containing Other Factors," *Journal of Quality Technology*, 27, 1, 13–33.
- Cornell, J. A. (2000), "Fitting a Slack-Variable Model to Mixture Data: Some Questions Raised," *Journal of Quality Technology*, 31, 133–147.
- Cornell, J. A. (2002), *Experiments with Mixtures, Designs, Models, and the Analysis of Mixture Data*, 3rd edition, John Wiley & Sons, New York.
- Covey-Crump, P. A. K., and Silvey, S. D. (1970), "Optimal Regression Designs with Previous Observations," *Biometrika*, 57, 551–566.
- Cox, D. R. (1958), *Planning of Experiments*, John Wiley & Sons, New York.
- Cox, D. R. (1971), "A Note on Polynomial Response Functions for Mixtures," *Biometrika*, 58, 155–159.
- Crosier, R. B. (1984), "Mixture Experiments: Geometry and Pseudo-components," *Technometrics*, 26, 209–216.
- Daniel, C. (1959), "Use of Half-Normal Plots in Interpreting Factorial Two-Level Experiments," *Technometrics*, 1, 311–342.
- Daniel, C. (1976), *Applications of Statistics to Industrial Experimentation*, John Wiley & Sons, Inc., New York.
- Del Castillo, E. (1996), "Multiresponse Process Optimization via Constrained Confidence Regions," *Journal of Quality Technology*, 28, 61–70.
- Del Castillo, E., Fan, S.-K. S., and Semple, J. (1997), "Computation of Global Optima in Dual Response Systems," *Journal of Quality Technology*, 29, 347–353.
- Del Castillo, E., and Montgomery, D. C. (1993), "A Nonlinear Programming Solution to the Dual Response Problem," *Journal of Quality Technology*, 25, 3.
- Del Castillo, E., Montgomery, D. C., and McCarville, D. R. (1996), "Modified Desirability Functions for Multiple Response Optimization," *Journal of Quality Technology*, 28, 347–356.
- Derringer, G., and Suich, R. (1980), "Simultaneous Optimization of Several Response Variables," *Journal of Quality Technology*, 12, 214–219.
- Donohue, J. M. (1994), "Experimental Designs for Simulation," *Winter Simulation Conference*, 200–206.

- Drain, D., Carlyle, W. M., Montgomery, D. C., Borror, C. M., and Anderson-Cook, C. M. (2004), "A Genetic Algorithm Hybrid for Constructing Optimal Response Surface Designs," *Quality and Reliability Engineering International*, 20, 637–650.
- Draper, N. R. (1963), "Ridge Analysis of Response Surfaces," *Technometrics*, 5, 469–479.
- Draper, N. R. (1982), "Center Points in Second-Order Response Surface Designs," *Technometrics*, 24, 127–133.
- Draper, N. R. (1985), "Small Composite Designs," *Technometrics*, 27, 173–180.
- Draper, N., and Sanders, E. (1988), "Designs for Minimizing Bias Estimation," *Technometrics*, 30, 319–324.
- Dykstra, O., Jr. (1966), "The Orthogonalization of Undesigned Experiments," *Technometrics*, 8, 279–290.
- Edwards, D. J., and Mee, R. W. (2008), "Empirically Determined  $p$ -Values for Length  $t$ -Statistics," *Journal of Quality Technology*, 40(4), 368–380.
- Ellerton, R. R. W. (1978), "Is the Regression Equation Adequate—A Generalization," *Technometrics*, 20, 313–316.
- Engle, J. (1992), "Modeling Variation in Industrial Experiments," *Journal of the Royal Statistical Society Series C (Applied Statistics)*, 41, 579–593.
- Engle, J., and Huele, A. F. (1996), "A Generalized Linear Modeling Approach to Robust Design," *Technometrics*, 38, 365–373.
- Fan, S.-K. S. (2000), "A Generalized Global Optimization Algorithm for Dual Response Systems," *Journal of Quality Technology*, 32, 444–456.
- Fang, K. T. (1980), "The Uniform Design: Application of Number—Theoretic Methods in Experimental Design," *Acta Mathematicae Applicatae Sinica*, 3, 363–372.
- Fang, K. T., Li, R., and Sudjianto, A. (2006), *Design and Modeling for Computer Experiments*, Taylor & Francis Group, Boca Raton, FL.
- Giovannitti-Jensen, A., and Myers, R. H. (1989), "Graphical Assessment of the Prediction Capability of Response Surface Designs," *Technometrics*, 31, 159–171.
- Goldfarb, H. B., Borror, C. M., and Montgomery, D. C. (2003), "Mixture-Process Variable Experiments with Noise Variables," *Journal of Quality Technology*, 35, 393–405.
- Goldfarb, H. B., Anderson-Cook, C. M., Borror, C. M., and Montgomery, D. C. (2004a), "Fraction of Design Space to Assess the Prediction Capability of Mixture and Mixture-Process Designs," *Journal of Quality Technology*, 36, 169–179.
- Goldfarb, H. B., Borror, C. M., Montgomery, D. C., and Anderson-Cook, C. M. (2004b), "Three-Dimensional Variance Dispersion Graphs for Mixture-Process Experiments," *Journal of Quality Technology*, 36, 109–124.
- Goldfarb, H. B., Borror, C. M., Montgomery, D. C., and Anderson-Cook, C. M. (2004c), "Evaluation of Statistical Designs for Mixture-Process Variable Experiments with Noise Variables," *Journal of Quality Technology*, 36, 245–262.
- Goldfarb, H. B., Borror, C. M., Montgomery, D. C., and Anderson-Cook, C. M. (2005), "Using Genetic Algorithms to Generate Mixture-Process Experimental Designs Involving Control and Noise Variables," *Journal of Quality Technology*, 37, 60–74.
- Goos, P. (2002), *The Optimal Design of Blocked and Split-Plot Experiments*, Springer, New York.
- Goos, P., and Donev, A. N. (2007), "Tailor-Made Split-Plot Designs for Mixture and Process Variables," *Journal of Quality Technology*, 39, 326–339.
- Goos, P., and Vanderbroek, M. (2004), "Outperforming Completely Randomized Designs," *Journal of Quality Technology*, 36, 12–26.
- Graybill, F. (1976), *Introduction to the Theory of Linear Statistical Models*, Duxbury Press, Boston.

- Grego, J. M. (1993), "Generalized Linear Models and Process Variation," *Journal of Quality Technology*, 25, 288–295.
- Gunst, R. F., and Mason, R. L. (1979), "Some Considerations in the Evaluation of Alternative Prediction Equations," *Technometrics*, 21, 55–63.
- Hackney, H., and Jones, P. R. (1969), "Response Surface for Dry Modulus of Rupture and Drying Shrinkage," *American Ceramic Society Bulletin*, 46.
- Hall, M., Jr. (1961), Hadamard matrix of order 16, *Jet Propulsion Laboratory Research Summary*, No. 1, pp. 21–26.
- Hamada, M., and Nelder, J. A. (1997), "Generalized Linear Models for Quality-Improvement Experiments," *Journal of Quality Technology*, 29, 292–304.
- Hamada, M., and Balakrishnan, N. (1998), "Analyzing Unreplicated Factorial Experiments: A Review with Some New Proposals" (with Discussion), *Statistica Sinica*, 8, 1–41.
- Hartley, H. O. (1959), "Smallest Composite Design for Quadratic Response Surfaces," *Biometrics*, 15, 611–624.
- Harvey, A. C. (1976), "Estimation of Regression Models with Multiplicative Heteroscedasticity," *Econometrica*, 44, 461–475.
- Haykin, S. (1994), *Neural Networks: A Comprehensive Foundation*, Macmillan, New York.
- Hebble, T. L., and Mitchell, T. J. (1972), "Repairing Response Surface Designs," *Technometrics*, 14, 767–779.
- Heinsmann, J. A., and Montgomery, D. C. (1995), "Optimization of a Household Product Formulation Using a Mixture Experiment," *Quality Engineering*, 7, 583–600.
- Heredia-Langner, A., Carlyle, W. M., Montgomery, D. C., Borror, C. M., and Runger, G. C. (2003), "Genetic algorithms for the construction of  $D$ -optimal designs," *Journal of Quality Technology*, 35, 28–46.
- Heredia-Langner, A., Carlyle, W. M., Montgomery, D. C., and Borror, C. M. (2004), "Model-Robust Optimal Designs: A Genetic Algorithm Approach," *Journal of Quality Technology*, 36, 263–279.
- Higdon, D., Gattiker, J., Williams, B., and Rightley, M. (2008), "Computer Model Calibration Using High-Dimensional Output," *Journal of the American Statistical Association*, 103, 570–583.
- Higdon, D., Kennedy, M., Cavendish, J. C., Cafeo, J. A., and Ryne, R. D. (2004), "Combining Field Data and Computer Simulations for Calibration and Prediction," *SIAM Journal on Scientific Computing*, 26, 448–466.
- Hill, R. C., Judge, G. G., and Fomby, T. B. (1978), "Testing the Adequacy of a Regression Model," *Technometrics*, 20, 491–494.
- Hill, R. W. (1979), "On Estimating the Covariance Matrix of Robust Regression  $M$ -Estimates," *Communications in Statistics, Series A*, 8, 1183–1196.
- Hill, W. J., and Hunter, W. G. (1996), "A Review of Response Surface Methodology: A Literature Review," *Technometrics*, 8, 571–590.
- Himmel, C. D., and May, G. S. (1991), "Advantages of Plasma Etch Modeling Using Neural Networks over Statistical Techniques," *IEEE Transactions on Semiconductor Manufacturing*, 6, 103–111.
- Hoerl, A. E. (1959), "Optimum Solution of Many Variables Equations," *Chemical Engineering Progress*, 55, 67–78.
- Hoerl, A. E. (1964), "Ridge Analysis," *Chemical Engineering Symposium Series*, 60, 67–77.
- Hoerl, R. W. (1985), "Ridge Analysis 25 Years Later," *The American Statistician*, 39, 186–193.
- Hoke, A. T. (1974), "Economical Second-Order Designs Based on Irregular Fractions of the  $3^n$  Factorial," *Technometrics*, 17, 375–384.
- Hussain, A. S., Yu, X. Q., and Johnson, R. D. (1991), "Application of Neural Computing in Pharmaceutical Product Development," *Pharmaceutical Research*, 8, 1248–1252.

- Johnson, M. E., Moore, L. M., and Ylvisaker, D. (1990), "Minimax and Maxmin Distance Design," *Journal of Statistical Planning and Inference*, 26, 131–148.
- Johnson, M. E., and Jones, B. (2010), "Classical Design Structure of Orthogonal Designs with Six to Eight Factors and Sixteen Runs," *Quality and Reliability Engineering International*, 27, 61–70.
- Johnson, R. T., Montgomery, D. C., and Jones, B. (2011), "An Empirical Study of the Prediction Performance of Space-Filling Designs," *International Journal of Experimental Design and Process Optimisation*, 2, 1–18.
- Johnson, R. T., Montgomery, D. C., Jones, B., and Parker, P. A. (2010), "Comparing Computer Experiments for Fitting High-Order Polynomial Models," *Journal of Quality Technology*, 42(1), 86–102.
- Jones, B., and Johnson, R. T. (2009), "The Design and Analysis of the Gaussian Process Model," *Quality and Reliability Engineering International*, 25, 515–524.
- Jones, B., and Montgomery, D. C. (2010), "Alternatives to Resolution IV Screening Designs in 16 Runs," *International Journal of Experimental Design and Process Optimisation*, 1(4), 285–295.
- Jones, B., and Nachtsheim, C. J. (2011a), "A Class of Three-Level Designs for Definitive Screening in the Presence of Second-Order Effects," *Journal of Quality Technology*, 43(1), 1–15.
- Jones, B., and Nachtsheim, C. J. (2011b), "Efficient Designs with Minimal Aliasing," *Technometrics*, 53, 62–71.
- Jones, B., and Nachtsheim, C. J. (2013), "Definitive Screening Designs with Added Two-Level Categorical Factors," *Journal of Quality Technology*, 45(2), 121–129.
- Jones, B., and Nachtsheim, C. J. (2015), "Blocking Schemes for Definitive Screening Designs," *Technometrics* (to appear).
- Jones, B., Shinde, S. M., and Montgomery, D. C. (2015), "Alternatives to Resolution III Regular Fractional Factorial Screening Designs for 9–14 Factors in 16 Runs," *Applied Stochastic Models in Business and Industry*, 31, 50–58.
- Joseph, V. R. and Yan, H. (2014), "Engineering-Driven Statistical Adjustment," *Technometrics* (to appear).
- Karson, M. J., Manson, H. R., and Hader, R. J. (1969), "Minimum Bias Estimation and Experimental Design for Response Surfaces," *Technometrics*, 11, 461–475.
- Kennedy, M. C. and O'Hagan, A. (2001), "Bayesian Calibration of Computer Models," *Journal of the Royal Statistical Society, Ser. B.*, 63(3), 425–464.
- Khattree, R. (1996), "Robust Parameter Design: A Response Surface Approach," *Journal of Quality Technology*, 28, 187–198.
- Khuri, A. I. (editor) (2006), *Response Surface Methodology And Related Topics*. World Scientific Publishing Co.
- Khuri, A. I., and Conlon, M. (1981), "Simultaneous Optimization of Multiple Responses Represented by Polynomial Regression Functions," *Technometrics*, 23, 363–375.
- Khuri, A. I., and Cornell, J. A. (1996), *Response Surfaces*, 2nd edition, Marcel Dekker, New York.
- Khuri, A. I., Harrison, J. M., and Cornell, J. A. (1999), "Using Quantile Plots of the Prediction Variance for Comparing Designs for a Constrained Mixture Region: An Application Involving a Fertilizer Experiment," *Journal of the Royal Statistical Society Series C (Applied Statistics)*, 48, 1–12.
- Khuri, A. I., Kim, H. J., and Um, Y. (1996), "Quantile Plots of the Prediction Variance for Response Surface Designs," *Computational Statistics and Data Analysis*, 22, 395–407.
- Kiefer, J. (1959), "Optimum Experimental Designs," *Journal of the Royal Statistical Society, Series B*, 21, 272–304.
- Kiefer, J. (1961), "Optimum Designs in Regression Problems. II," *Annals of Mathematical Statistics*, 32, 298–325.

- Kiefer, J., and Wolfowitz, J. (1959), "Optimum Designs in Regression Problems," *Annals of Mathematical Statistics*, 30, 271–294.
- Koons, G. F., and Wilt, M. H. (1985), "Design and Analysis of an ABS Pipe Compound Experiment," *Experiments in Industry: Design, Analysis, and Interpretation of Results*, edited by R. D. Snee, L. B. Hare, and R. Trout, American Society in Quality Control, Milwaukee, 111–117.
- Kowalski, S. M., Borror, C. M., and Montgomery, D. C. (2005), "A Modified Path of Steepest Ascent for Split-Plot Experiments," *Journal of Quality Technology*, 37, 75–83.
- Kowalski, S. M., Cornell, J. A., and Vining, G. G. (2002), "Split-Plot Designs and Estimation Methods for Mixture Experiments with Process Variables," *Technometrics*, 44, 72–79.
- Larntz, K., and Whitcomb, P. (1998). "Use of Replication in Almost Unreplicated Factorials." *Presented at the Fall Technical Conference*, Corning, NY.
- Lee, Y., and Nelder, J. A. (1998), "Letter to the Editor: Joint Modeling of Mean and Dispersion," *Technometrics*, 40, 168–175.
- Lenth, R. V. (1989). "Quick and Easy Analysis of Unreplicated Factorials." *Technometrics*, 31, 469–473.
- Lentner, M., and Bishop, T. (1993), *Experimental Design and Analysis*, Valley Book Company, Blacksburg, VA.
- Letsinger, J. D., Myers, R. H., and Lentner, M. (1996), "Response Surface Methods for Bi-randomization Structures," *Journal of Quality Technology*, 26, 381–397.
- Lewis, S. L. (1998), "Analysis of Designed Experiments Using Generalized Linear Models," Ph.D. Dissertation, Department of Industrial Engineering, Arizona State University (1998).
- Lewis, S. L., Montgomery, D. C., and Myers, R. H. (2001a), "Examples of Designed Experiments with Nonnormal Responses," *Journal of Quality Technology*, 33, 265–278.
- Lewis, S. L., Montgomery, D. C., and Myers, R. H. (2001b), "Confidence Interval Coverage and Precision for Designed Experiments Analyzed with Generalized Linear Models," *Journal of Quality Technology*, 33, 279–292.
- Li, J., Liang, L., Borror, C. M., Anderson-Cook, C. M., and Montgomery, D. C. (2009), "Graphical Summaries to Compare Prediction Variance Performance for Variations of the Central Composite Design for 6 to 10 Factors," *Quality Technology and Quantitative Management*, 6(4), 433–449.
- Liang, L., Anderson-Cook, C. M., and Robinson, T. J. (2006a), "Fraction of Design Space Plots for Split-Plot Designs," *Quality and Reliability Engineering International*, 22, 275–289.
- Liang, L., Anderson-Cook, C. M., Robinson, T. J., and Myers, R. H. (2006b), "Three-Dimensional Variance Dispersion Graphs For Split-Plot Designs," *Journal of Computational and Graphical Statistics*, 15, 757–778.
- Liang, L., Anderson-Cook, C. M., and Robinson, T. J. (2007), "Cost Penalized Estimation and Prediction Evaluation for Split-Plot Designs," *Quality and Reliability Engineering International*, 23, 577–596.
- Lin, C. F. D., Anderson-Cook, C. M., Hamada, M. S., Moore, L. M., and Sitter, R. R. (2015), "Using Genetic Algorithms to Design Experiments: A Review," *Quality and Reliability Engineering International*, 31, 155–167.
- Lin, D. K. J., and Tu, W. (1995), "Dual Response Surface Optimization," *Journal of Quality Technology*, 27, 248–260.
- Lind, E. E., Goldin, J., and Hickman, J. B. (1960), "Fitting Yield and Cost Response Surfaces," *Chemical Engineering Progress*, 56, 62.
- Loughin, T. M. (1998), "Calibration of the Lenth Test for Unreplicated Factorial Designs," *Journal of Quality Technology*, 30, 171–175.
- Loughin, T. M., and Noble, W. (1997), "A Permutation Test for Effects in an Unreplicated Factorial Design," *Technometrics*, 39, 180–190.



- Lu L., and Anderson-Cook, C. M. (2012), "Rethinking the Optimal Response Surface Design for a First-Order Model with Two-Factor Interactions, when Protecting against Curvature," *Quality Engineering*, 24 404–422.
- Lu, L., and Anderson-Cook, C. M. (2013), "Adapting the Hypervolume Quality Indicator to Quantify Trade-Offs and Search Efficiency for Multiple Criteria Decision-Making Using Pareto Fronts," *Quality and Reliability Engineering International*, 29, 1117–1133.
- Lu, L., and Anderson-Cook, C. M. (2014), "Balancing Multiple Criteria Incorporating Cost using Pareto Front Optimization for Split-Plot Designed Experiments," *Quality and Reliability Engineering International*, 30, 37–55.
- Lu, L., Anderson-Cook, C. M., and Robinson, T. J. (2011), "Optimization of Designed Experiments Based on Multiple Criteria Utilizing a Pareto Frontier," *Technometrics*, 53, 353–365.
- Lu, L., Anderson-Cook, C. M., and Robinson, T. J. (2012), "A Case Study to Demonstrate Pareto Frontiers for Selecting a Best Response Surface Design with Simultaneously Optimizing Multiple Criteria," *Applied Stochastic Models in Business and Industry*, 28, 206–221.
- Lu, L., Chapman, J., and Anderson-Cook, C. M. (2013), "A Case Study on Selecting A Best Allocation of New Data for Improving the Estimation Precision of System and Sub-System Reliability using Pareto Fronts," *Technometrics*, 55, 473–487.
- Lu, L., Johnson, M. E., and Anderson-Cook, C. M. (2014), "Selecting a Best Two Level 16-run Screening Design from the Catalog of Non-Isomorphic Regular and Non-Regular Designs for Six to Eight Factors," *Quality Engineering*, 26, 269–284.
- Lu, L., Robinson, T. J., and Anderson-Cook, C. M. (2014), "A Case Study to Select an Optimal Split-Plot Design for a Mixture-Process Experiment Based on Multiple Objectives," *Quality Engineering*, 26, 424–439.
- Lu., L., Anderson-Cook, C. M., and Lin, D. (2014) "Optimal Designed Experiments Using a Pareto Front Search for Focused Preference of Multiple Objectives," *Computational Statistics and Data Analysis*, 71, 1178–1192.
- Lucas, J. M. (1974), "Optimum Composite Designs," *Technometrics*, 16, 561–567.
- Lucas, J. M. (1976), "Which Response Surface Design Is Best," *Technometrics*, 18, 411–417.
- Lucas, J. M. (1989), "Achieving a Robust Process Using Response Surface Methodology," *American Statistical Association Proceedings Sesquicentennial Invited Paper Sessions 1988 and 1989*, 579–593.
- Lucas, J. M. (1994), "How to Achieve a Robust Process Using Response Surface Methodology," *Journal of Quality Technology*, 26, 248–260.
- Lucas, J. M., and Ju, H. L. (1992), "Split Plotting and Randomization in Industrial Experiments," *ASQC Quality Transactions*, Nashville, TN.
- Marquardt, D. M., and Snee, R. D. (1975), "Ridge Regression in Practice," *The American Statistician*, 29, 3–20.
- Martin, B., Parker, D., and Zenick, L. (1987), "Minimize Slugging by Optimizing Controllable Factors on Topaz Windshield Molding," *Fifth Symposium on Taguchi Methods*, American Supplier Institute, Inc., Dearborn, MI, 519–526.
- Martins, S. A. M., Prazeres, D. M. F., Fonseca, L. P., and Monteiro, G. A. (2009), "Application of Central Composite Design for DNA Hybridization onto Magnetic Microparticles," *Analytical Biochemistry*, 391, 17–23.
- Maruyama, S. A., Palombinin, S. V., Claus, T., Carbonera, F., Montanher, P. F., de Souza, N. E., Visertainer, J. V., Gomes, S. T. M., and Matsushita, M. (2013), "Application of Box–Behnken Design to the Study of Fatty Acids and Antioxidant Activity from Enriched White Bread," *Journal of Brazilian Chemical Society*, 24, 1520–1529.
- McCullagh, P., and Nelder, J. A (1989), *Generalized Linear Models*, 2nd edition, Chapman and Hall, New York.

- McKay, M. D., Beckman, R. J., and Conover, W. J. (1979), "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," *Technometrics*, 21, 239–245.
- McLean, R. A., and Anderson, V. L. (1966), "Extreme Vertices Design of Mixture Experiments," *Technometrics*, 8, 447–454.
- Mead, R., and Pike, D. J. (1975), "A Review of Response Surface Methodology from a Biometric Viewpoint," *Biometrics*, 31, 803–851.
- Meyers, R. K., and Nachtsheim, C. J. (1988), "Constructing Exact  $D$ -Optimal Experimental Designs by Simulated Annealing," *American Journal of Mathematical and Management Sciences*, 8, 329–359.
- Mitchell, T. J., and Bayne, C. K. (1978), " $D$ -Optimal Fractions of Three-Level Factorial Designs," *Technometrics*, 20, 369–380.
- Mitchell, T. J., and Morris, M. D. (1992), "The Spatial Correlation Approach to Response Surface Estimation," *Winter Simulation Conference*, 565–571.
- Modi, O. P., Yadav, R. P., Mondal, D. P., Dasgupta R., Das, S., and Yegneswaran, A. H. (2001), "Abrasive Wear Behaviour of Zinc-Aluminum Alloy–10%  $Al_2O_3$  Composite through Factorial Design of Experiment," *Journal of Materials Science*, 36, 1601–1607.
- Montgomery, D. C. (1991), "Using Fractional Factorial Designs for Robust Process Development," *Quality Engineering*, 3, 193–205.
- Montgomery, D. C. (1999), "Experimental Design for Product and Process Design and Development" (with comment), *Journal of the Royal Statistical Society Series D (The Statistician)*, 38, 159–177.
- Montgomery, D. C. (2013), *Design and Analysis of Experiments*, 7th edition, John Wiley & Sons, New York.
- Montgomery, D. C., Peck, E. A., and Vining, G. G. (2012), *Introduction to Linear Regression Analysis*, 5th edition, John Wiley & Sons, New York.
- Montgomery, D. C., and Runger, G. C. (1996), "Foldovers of  $2^{k-p}$  Resolution IV Designs," *Journal of Quality Technology*, 28, 446–450.
- Montgomery, D. C., and Voth, S. R. (1994), "Multicollinearity and Leverage in Mixture Experiments," *Journal of Quality Technology*, 26, 96–108.
- Montgomery, D. C., and Weatherby, G. (1979), "Factor Screening Methods in Computer Simulation," *Winter Simulation Conference*, 347–358.
- Morris, M. D. (2000), "A Class of Three-Level Experimental Designs for Response Surface Modeling," *Technometrics*, 42, 111–121.
- Myers, R. H. (1976), *Response Surface Methodology*, published by author, Blacksburg, VA.
- Myers, R. H. (1990), *Classical and Modern Regression with Applications*, 2nd edition, Duxbury Press, Boston.
- Myers, R. H. (1999), "Response Surface Methodology: Current Status and Future Directions" (with discussion), *Journal of Quality Technology*, 31, 30–44.
- Myers, R. H., and Carter, W. H., Jr. (1973), "Response Surface Techniques for Dual Response Systems," *Technometrics*, 15, 301–317.
- Myers, R. H., Khuri, A. I., and Carter, W. H. (1989), "Response Surface Methodology: 1966–1988," *Technometrics*, 3, 137–157.
- Myers, R. H., Khuri, A. I., and Vining, G. (1992a), "Response Surface Alternatives to the Taguchi Robust Parameter Design Approach," *The American Statistician*, 46, 2, 131–139.
- Myers, R. H., Kim, Y., and Griffiths, K. (1997), "Response Surface Methods and Use of Noise Variables," *Journal of Quality Technology*, 29, 429–441.
- Myers, R. H., and Milton, S. (1991), *A First Course in the Theory of Linear Statistical Models*, Duxbury Press, Boston.

- Myers, R. H., and Montgomery, D. C. (1997), "A Tutorial on Generalized Linear Models," *Journal of Quality Technology*, 29, 274–291.
- Myers, R. H., Montgomery, D. C., Vining, G. G., and Robinson, T. J. (2011), *Generalized Linear Models with Applications in Engineering and the Sciences*, 2nd edition, John Wiley & Sons, Hoboken, NJ.
- Myers, R. H., Montgomery, D. C., Vining, G. G., Borror, C. M., and Kowalski, S. M. (2004), "Response Surface Methodology: A Retrospective and Literature Survey," *Journal of Quality Technology*, 36, 53–77.
- Myers, R. H., Vining, G., Giovannitti-Jensen, A., and Myers, S. L. (1992b), "Variance Dispersion Properties of Second-Order Response Surface Designs," *Journal of Quality Technology*, 24, 1–11.
- Nair, V. N., editor (1992), "Taguchi's Parameter Design: A Panel Discussion," *Technometrics*, 34, 2, 127–161.
- Nair, V. N., and Pregibon, D. (1986), "A Data Analysis Strategy for Quality Engineering Experiments," *AT&T Technical Journal*, 74–84.
- Nalimov, V. V., Golikova, T. I., and Mikeshina, N. G. (1970), "On Practical Use of the Concept of *D*-Optimality," *Technometrics*, 12, 799–812.
- Nelder, J. A., and Lee, Y. (1991), "Generalized Linear Models for the Analysis of Taguchi-Type Experiments," *Applied Stochastic Models and Data Analysis*, 7, 107–120.
- Notz, W. (1982), "Minimal Point Second Order Designs," *Journal of Statistical Planning and Inference*, 6, 47–58.
- Oehlert, G., and Whitcomb, P. (2002), "Small, Efficient Equireplicated Resolution V Fractions of  $2^k$  Designs and Their Application to Central Composite Designs," Paper presented at the 46<sup>th</sup> Annual Fall Technical Conference of the ASQ and ASA, Valley Forge, PA. Also available at <http://www.statease.com/>.
- Ozol-Godfrey, A., and Anderson-Cook C. M. (2007), "Fraction of Design Space Plots for Examining Mixture Design Robustness to Measurement Errors," *Journal of Statistics and Applications*, 1, 171–183.
- Ozol-Godfrey, A., Anderson-Cook, C. M., and Montgomery, D. C. (2005), "Fraction of Design Space Plots for Examining Model Robustness," *Journal of Quality Technology*, 37, 223–235.
- Ozol-Godfrey, A., Anderson-Cook, C. M., Robinson, T. J. (2008), "Fraction of Design Space Plots for Generalized Linear Models," *Journal of Statistical Planning and Inference*, 138, 203–219.
- Park, Y.-J., Richardson, D. E., Montgomery, D. C., Ozol-Godfrey, A., Borror, C. M., and Anderson-Cook, C. M. (2005), "Prediction Variance Properties of Second-Order Designs for Cuboidal Regions," *Journal of Quality Technology*, 37(4), 253–266.
- Parker, P. A., Anderson-Cook, C. M., and Robinson, T. J. (2008), "Robust Split-Plot Designs," *Quality and Reliability Engineering International*, 24, 107–121.
- Parker, P. A., Kowalski, S. M., and Vining, G. G. (2006), "Classes of Split-Plot Response Surface Designs for Equivalent Estimation," *Quality and Reliability Engineering International*, 22, 291–305.
- Parker, P. A., Kowalski, S. M., and Vining, G. G. (2007), "Construction of Balanced Equivalent Estimation Second-Order Split-Plot Designs," *Technometrics*, 49, 56–65.
- Piepel, G. F. (1982), "Measuring Component Effects in Constrained Mixture Experiments," *Technometrics*, 25, 97–101.
- Piepel, G. F. (1988), "Programs for Generating Extreme Vertices and Centroids of Linearly Constrained Experimental Regions," *Journal of Quality Technology*, 20, 125–139.
- Piepel, G. F., and Cornell, J. A. (1994), "Mixture Experiment Approaches: Examples, Discussion, and Recommendations," *Journal of Quality Technology*, 26, 177–196.
- Pigantiello, Jr., J. J. (1993), "Strategies for Robust Multiresponse Quality Engineering," *IIE Transactions*, 25, 5–15.

- Pignatiello, J., and Ramberg, J. (1991), "Top Ten Triumphs and Tragedies of Genichi Taguchi," *Quality Engineering*, 4(2), 211–225.
- Plackett, R. L., and Burman, J. P. (1946), "The Design of Optimum Multifactorial Experiments," *Biometrika*, 33, 305–325.
- Pukelsheim, F. (1995), *Optimum Design of Experiments*, Chapman & Hall, New York.
- Ripley, B. D. (1994a), "Statistical Ideas for Selecting Network Architecture," in *Neural Networks: Artificial Intelligence and Industrial Applications*, edited by B. Kappen and S. Gielen, Springer-Verlag, New York, 183–190.
- Ripley, B. D. (1994b), "Neural Networks and Related Methods for Classification," *Journal of the Royal Statistical Society Series B (Theory and Methods)*, 56, 409–456.
- Robinson, T. J., and Anderson-Cook, C. M. (2011), "A Closer Look at D-Optimality for Screening Designs," *Quality Engineering*, 23, 1–14.
- Rodriguez, M., Jones, B., Borrer, C. M., and Montgomery, D. C. (2010), "Generating and Assessing Exact G-Optimal Designs," *Journal of Quality Technology*, 42(1), 3–29.
- Roquemore, K. G. (1976), "Hybrid Designs for Quadratic Response Surfaces," *Technometrics*, 18, 419–423.
- Rousseeuw, P. J., and Leroy, A. M. (1987), *Robust Regression and Outlier Detection*, John Wiley & Sons, New York.
- Rozum, M. A. (1990), "Effective Design Augmentation for Prediction," Thesis, Virginia Tech.
- Rozum, M., and Myers, R. H. (1991), "Variance Dispersion Graphs for Cuboidal Regions," paper presented at ASA Meetings, Atlanta, GA.
- Sacks, J., Schiller, S. B., and Welch, W. J. (1989a), "Designs for Computer Experiments," *Statistical Science*, 4, 409–435.
- Sacks, J., and Welch, W. J. (1989), "Design and Analysis of Computer Experiments," *Statistical Science*, 4, 409–435.
- Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989b), "Design and Analysis of Computer Experiments," *Statistical Science*, 4, 409–423.
- Santner, T. J., Williams, B. J., and Notz, W. I. (2003), *The Design and Analysis of Computer Experiments*, Springer Series in Statistics, Springer-Verlag, New York.
- Scheffé, H. (1958), "Experiments with Mixtures," *Journal of the Royal Statistical Society, Series B*, 20, 344–366.
- Scheffé, H. (1959), "Reply to Mr. Quenouille's Comments About my Paper on Mixtures," *Journal of the Royal Statistical Society, Series B*, 23, 171–172.
- Scheffé, H. (1965), "The Simplex-Centroid Design for Experiments with Mixtures," *Journal of the Royal Statistical Society, B*, 25, 235–263.
- Schmidt, S. R., and Launsby, R. G. (1990), *Understanding Industrial Designed Experiments*, CQG Limited, Colorado Springs, CO.
- Searle, S. R., Casella, G., and McCulloch, C. E. (1992), *Variance Components*, John Wiley & Sons, New York.
- Shewry, M. C., and Wynn, H. P. (1987), "Maximum Entropy Sampling," *Journal of Applied Statistics*, 14, 165–170.
- Shinde, S. M., Montgomery, D. C., and Jones, B. (2014), "Projection Properties of No-Confounding Designs for Six, Seven, and Eight Factors in 16 Runs," *International Journal of Experimental Design and Process Optimization*, 4(1), 1–26.
- Shoemaker, A. C., Tsui, K. L., and Wu, C. F. J. (1991), "Economical Experimentation Methods for Robust Design," *Technometrics*, 33, 415–427.
- Simpson, T., and Peplinski, J. (1997), "On the Use of Statistics in Design and the Implications for Deterministic Computer Experiments," *ASME Design Engineering Technical Conference*, 1–12.

- Slaagame, R. R., and Barton, R. R. (1997), "Factorial Hypercube Designs for Spatial Correlation Regression," *Journal of Applied Statistics*, 24, 483–501.
- Snee, R. D., and Marquardt, D. W. (1974), "Extreme Vertices Designs for Linear Mixture Models," *Technometrics*, 16, 399–408.
- St. John, R. C. (1984), "Experiments with Mixtures: Ill-Conditioning and Ridge Regression," *Journal of Quality Technology*, 16, 81–96.
- St. John, R. C., and Draper, N. R. (1975), "D-Optimality for Regression Designs: A Review," *Technometrics*, 17, 15–23.
- Steiner, S. H., and Hamada, M. (1997), "Making Mixtures Robust to Noise and Mixing Measurement Errors," *Journal of Quality Technology*, 29, 441–450.
- Suich, R., and Derringer, G. C. (1977), "Is the Regression Equation Adequate—One Criterion," *Technometrics*, 19, 213–216.
- Taguchi, G. (1986), *Introduction to Quality Engineering*, Asian Productivity Organization, UNIPUB, White Plains, NY.
- Taguchi, G. (1987), *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Cost*, UNIPUB/Kraus International, White Plains, NY.
- Taguchi, G., and Wu, Y. (1980), *Introduction to Off-Line Quality Control*, Central Japan Quality Control Association, Nagoya, Japan.
- Trinca, L. A., and Gilmour, S. G. (1998), "Variance Dispersion Graphs for Comparing Blocked Response Surface Designs," *Journal of Quality Technology*, 30, 314–327.
- Vining, G. (1993), "A Computer Program for Generating Variance Dispersion Graphs," *Journal of Quality Technology*, 25, 45–58.
- Vining, G. G. (1998), "A Compromise Approach to Multiresponse Optimization," *Journal of Quality Technology*, 30, 309–313.
- Vining, G. G., and Bohn, L. (1998), "Response Surfaces for the Mean and the Process Variance Using a Nonparametric Approach," *Journal of Quality Technology*, 30, 282–291.
- Vining, G. G., Cornell, J. A., and Myers, R. H. (1993), "A Graphical Approach for Evaluating Mixture Designs," *Journal of the Royal Statistical Society Series C (Applied Statistics)*, 42, 127–138.
- Vining, G. G., Kowalski, S. M., and Montgomery, D. C. (2005), "Response Surface Designs Within a Split-Plot Structure," *Journal of Quality Technology*, 37, 115–129.
- Vining, G. G., and Myers, R. H. (1990), "Combining Taguchi and Response Surface Philosophies: A Dual Response Approach," *Journal of Quality Technology*, 22, 38–45.
- Vining, G. G., and Myers, R. H. (1991), "A Graphical Approach for Evaluating Response Surface Designs in Terms of the Mean Squared Error of Prediction," *Technometrics*, 33, 315–326.
- Vining, G. G., and Schaub, D. (1996), "Experimental Designs for Estimating Both Mean and Variance Functions," *Journal of Quality Technology*, 28, 135–147.
- Weaver, B. P., Warr, R. L., Anderson-Cook, C. M., and Higdon, D. M. (2015) "Visualizing Discrepancies from Nonlinear Models and Computer Experiments," *Statistical Analysis and Data Mining*, 8, 274–286.
- William, B., Higdon, D., Gattiker, J., Moore, L., McKay, M., and Keller-McNulty, S. (2006), "Combining Experimental Data and Computer Simulations, with an Application to Flyer Plate Experiments," *Bayesian Analysis*, 1, 765–792.
- Welch, W. J. (1982), "Branch-and-Bound Search for Experimental Designs Based on D Optimality and Other Criteria," *Technometrics*, 24, 41–48.
- Welch, W. J., Yu, T. K., Kang, S. M., and Sacks, J. (1990), "Computer Experiments for Quality Control by Parameter Design," *Journal of Quality Technology*, 22, 15–22.
- Welch, W. J., Buck, R. J., Sacks, J., Wynn, H. P., Mitchell, T. J., and Morris, M. D. (1992), "Screening, Predicting, and Computer Experiments," *Technometrics*, 34, 15–25.

- Welsch, R. E. (1975), "Confidence Regions for Robust Regression," *Proceedings of the Statistical Computing Section*, American Statistical Association, Washington, DC.
- Wolfinger, R., Tobias, R., and Sall, J. (1994), "Computing Gaussian Likelihood and their Derivatives for General Linear Mixed Models," *Siam Journal on Scientific Computation*, 15, 1294–1310.
- Yates, F. (1935), "Complex Experiments," *Journal of the Royal Statistical Society*, Supplement 2, 181–247.
- Yates, F. (1937), "*Design and Analysis of Factorial Experiments*," Technical Communication No. 35, Imperial Bureau of Soil Sciences, London.
- Xiao, L., Lin, D. K. J., and Bai, F. (2012) "Construction of Definitive Screening Designs Using Conference Matrices," *Journal of Quality Technology*, 44(1), 2–8.
- Ye, K., and Hamada, M. (2000), "Critical Values of the Lenth Method for Unreplicated Factorial Designs," *Journal of Quality Technology*, 32, 57–66.
- Zahran, A., Anderson-Cook, C. M., and Myers, R. H. (2003), "Fraction of Design Space to Assess the Prediction Capability of Response Surface Designs," *Journal of Quality Technology*, 35, 377–386.

# INDEX

- $2^2$  factorial design, 82
- $2^3$  factorial design, 94
- $2^k$  factorial design, 81, 103
- $2^{k-1}$  fractional factorial design, 162
- $2^{k-2}$  fractional factorial design, 174
- $2^{k-p}$  fractional factorial design, 184
  
- Achieving specific customer requirements, 20
- Active lower and upper bounds in mixture experiments, 747
- Additive desirability function, 329
- Adjusted deviance residuals, 606
- Adjusted  $R^2$ , 26
- Algorithms for constructing
  - computer-generated designs, 494
- Alias matrix, 199, 374
- Alias matrix in fractional factorial designs, 199
- Aliasing in a fractional factorial design, 163, 195
- Alternate fraction, 164
- Analysis of nonregular designs, 213
- Analysis of variance, 22, 84, 85
- Antagonistic blending, 699
- A-optimal design, 470
- Augmented simplex mixture designs, 707
- Augmenting definitive screening designs, 422
- Average prediction variance, 548
  
- Average squared bias, 548
- Average squared error loss design criterion, 548, 558
- Axial check blends, 708
- Axial runs, 48, 134, 281, 394, 482
- Axial runs in mixture experiments, 707, 708
  
- Balanced incomplete block design, 413
- Balanced split-plot design, 588
- Binary mixtures, 694
- Blocking a replicated  $2^k$  factorial, 135
- Blocking and confounding, 137
- Blocking in the  $2^k$  factorial, 135
- Blocking second-order designs, 422
- Boundary point design, 704, 707
- Box-Behnken designs, 413, 415
- Box-Cox method for choosing a transformation, 65
  
- Canonical analysis of a second-order model, 278
- Canonical link function, 593
- Canonical mixture model, 698
- Cell plot of the alias structure in a fractional factorial design, 202
- Center points in the  $2^k$  factorial design, 130, 134

- Center runs for the rotatable CCD, 403, 480  
 Central composite design, 49, 81, 134, 393  
 Choice between cuboidal and spherical design regions, 411  
 Choice of axial values in the CCD, 482  
 Coded variables for a  $2^k$  factorial, 86  
 Combined array design, 635, 661  
 Combining information from physical and computer experiments, 574  
 Complete mixtures, 694  
 Completely randomized design, 141  
 Component axis, 707  
 Concurrent engineering, 11  
 Conditions for orthogonal blocking in second-order designs, 424  
 Conference matrix, 387  
 Confidence interval for direction of steepest ascent, 247  
 Confidence interval on a regression coefficient, 32  
 Confidence interval on the mean response, 33  
 Confidence intervals on effects in a  $2^k$  factorial, 103  
 Confidence intervals on eigenvalues in canonical analysis, 304  
 Confidence region on the location of the stationary point, 299  
 Confirmation experiment, 239, 628  
 Constrained optimization for handling multiple responses, 340  
 Constraints in mixture experiments, 731, 732, 743, 747, 758  
 Constructing a  $2^{k-1}$  fractional factorial design, 165  
 Contour plots in analyzing second-order response surfaces, 286  
 Contrast, 84  
 Control factors in robust design, 619  
 Control-by-noise interactions in robust design, 631  
 Cook's distance, 43  
 Coordinate exchange, 494  
 Correlation matrix, 202  
 Covariance matrix of the least squares estimators, 22  
 Cox's direction, 784  
 Crossed array designs, 622, 625, 631, 657  
 Cuboidal design region, 409  
  
*D*-efficient design, 130, 469  
 Defining relation for a fractional factorial design, 163, 178  
 Definitive screening designs, 385, 417, 422  
 Definitive screening designs with two-level qualitative variables, 516  
 Derived power vectors, 797  
 Design matrix for a  $2^k$  factorial, 94  
 Design moments, 399  
 Design projection in a  $2^k$  factorial, 114  
 Design projection in a fractional factorial design, 166, 187  
 Design resolution, 164  
 Designs for first-order models, 375  
 Designs for second-order models, 393  
 Desirability functions, 328, 341  
 Desirable properties of a design, 369, 370  
 Deterministic computer models, 563  
 Deviance residuals, 605  
 Direct variance modeling, 655  
 Dispersion effects, 622, 673  
 Dispersion effects in fractionated designs, 672, 682  
 Distance-based designs for mixture experiments, 750  
 DMRCs approach for sequential optimization, 350  
*D*-optimal designs, 126, 467  
 Double linear regression method, 305  
 Dual response approach in robust design, 635, 639, 640  
 Dual response optimization method, 349  
  
 Effect of center runs, 380  
 Effect sparsity, 386  
 Empirical model, 13  
 Equiradial designs, 463, 803  
 Equivalent estimation split-plot design, 589  
 Errors in design variables, 560  
 Estimated prediction variance, 407  
 Estimation of  $\sigma^2$ , 22  
 Experimenter, 2  
 Experiments with computer models, 563, 567, 569  
 Exponential family of distributions, 591  
 Externally studentized residual, 41  
 Extra sum of squares method, 28  
  
 Face-centered cube, 408  
 Fast flexible filling designs, 569  
 Finding inactive components, 784  
 First-order model, 3  
 First-order model with interaction, 4  
 First-order orthogonal design, 376  
 Fold-over of fractional factorials, 190, 193, 194



- Fraction of design space plot, 128, 479
- Fractional factorial designs, 161
- Full model, 29
  
- Gaussian process IMSE designs, 569
- Gaussian process model, 570
- $G$ -efficient design, 130, 471
- Generalized linear models, 591, 657, 658
- Generating relations in a fractional factorial design, 174
- Geometric notation for a  $2^k$  factorial, 94
- $G$ -optimal designs, 127, 470
  
- Half-normal probability plot of effects in a  $2^k$  factorial, 109
- Hat matrix in regression, 38, 42
- Heredity, 213
- Hexagon design, 463
- Hierarchy principle, 215
- Hoke designs, 452
- Hybrid designs, 455, 484
  
- Ill-conditioning in mixture experiments, 734
- Independent variables, 2
- Indicator variables, 57
- Influence and leverage in regression, 42
- Influence on regression coefficients, 43
- Input variables, 2
- Interaction and curvature considerations in steepest ascent, 241
- Interaction terms, 14
- $I$ -optimal designs, 12, 470
- Iteratively reweighted least squares, 594
  
- Joint confidence region on regression coefficients, 33
  
- Khuri-Conlon method, 350
- Koshal designs, 454
  
- Lack of fit, 91, 132
- Latin hypercube designs, 568
- Least squares estimators, 15, 17, 22, 87
- Least squares normal equations, 16
- Lenth's method, 113
- Leverage points, 42
- Linear blending, 699
- Linear regression analysis, 6, 13
- Link function, 592
- Location effects, 622, 673
- Log link, 592
- Logistic regression model, 593
  
- Logit link, 592
- Log-linear model of process variance, 655, 680
- Lower bound constraints in mixture experiments, 732
- Low-order polynomial approximating function, 3
- $L$ -pseudocomponents, 733
  
- Main effect of a factor, 83
- Main effects model, 4
- Mapping a response surface, 9
- Maximin Latin hypercube designs, 568
- Maximum entropy designs, 560
- Maximum likelihood estimation, 593
- Maximum potential designs, 569
- Mean square error, 547
- Metamodels, 563
- Method of least squares, 6, 14, 15
- Minimal resolution IV designs, 197
- Minimax Latin hypercube designs, 568
- Minimum bias design, 551, 553, 554, 556
- Minimum bias estimation, 575
- Minimum-run resolution V designs, 412
- Mixed resolution designs, 662, 665
- Mixture experiment, 11, 693
- Mixture plot, 345
- Mixture-process model, 763
- Model bias, 373, 543, 545
- Model dependent estimate of  $\sigma^2$ , 23
- Model deviance in a GLM, 595
- Model independent estimate of  $\sigma^2$ , 23, 45
- Moment generating function, 800
- Model matrix, 16
- Moment matrix, 400, 402, 467, 550
- Moment matrix for a rotatable design, 402, 797
- Multicollinearity in mixture experiments, 734
- Multicomponent constraints, 758
- Multiple regression models, 13, 14
- Multiple responses, 325, 338
- Multiplicative desirability function, 329
  
- Neural networks, 579
- No-confounding designs, 207, 209
- Noise factors, 12, 620
- Nonisomorphic designs, 203
- Nonlinear blending, 699
- Nonlinear programming methods, 340
- Nonregular design, 195, 200, 203, 209
- Normal probability plot of effects in a  $2^k$  factorial, 108
- Normal probability plot of residuals, 37

- Odds ratio, 601
- Operability region, 8, 371
- Optimal design augmentation, 201
- Optimal designs, 125
- Orthogonal array, 371
- Orthogonal blocking in second-order designs, 422
- Orthogonal blocking in the Box-Behnken design, 428
- Orthogonal blocking in the CCD, 425
- Orthogonal blocking with definitive screening designs, 429
- Orthogonal design, 97, 372, 378, 381
- Orthogonal main effect plans, 371
- Overdispersion, 597
- Overlaying contour plots, 329
  
- Pareto front optimization, 327, 343
- Partial aliasing, 195
- Partial fold-over of a fractional factorial, 201
- Partial  $F$ -test, 30
- Partial regression coefficients, 13
- Path of steepest ascent, 234, 237
- Pearson  $\chi^2$  goodness-of-fit, 605
- Pearson  $\chi^2$  residuals, 598
- Phase one of RSM, 7
- Phase two of RSM, 7
- Phase zero of RSM, 7
- Pick-the-winner optimization in robust design, 625
- Piepel's direction, 785
- Plackett-Burman designs, 195
- Plot of residuals versus predicted response, 37
- Point exchange, 494
- Poisson regression model, 593
- Power family transformations, 65
- Practical guidance regarding steepest ascent, 242
- Prediction interval on future observations in regression, 35
- Prediction of new observations in regression, 35
- Prediction performance of combined array designs, 665
- Prediction profiler in JMP, 55, 104, 287
- Prediction variance contour plots, 477
- Prediction variance for first-order models, 389
- Prediction variance profiler in JMP, 128
- Predictor variables, 13
- PRESS residuals, 39
- PRESS statistic, 39
- Principal fraction, 163, 174
  
- Process optimization, 233
- Process robustness studies, 620
- Process variables in mixture experiments, 696, 763
- Product formulation and mixture experiments, 11, 695
- Product formulation, 11
- Propagation of error, 637
- Properties of the least squares estimators, 22
- Pure blends, 694
- Pure error, 91
- Pure quadratic curvature, 132
  
- Qualitative variables in regression, 55
- Qualitative variables in RSM, 508, 516
  
- $R^2$ , 26
- Randomized blocks, 141
- Range of factors in steepest ascent, 245
- Ratios of components in mixture experiments, 759
- Reciprocal link, 592, 592
- Reduced model, 29, 91
- Reference mixture, 716
- Region moment matrices, 551
- Region moments, 473, 550
- Region of experimentation, 8
- Region of interest, 8, 371
- Regression coefficients, 13, 14
- Regression model for a  $2^k$  factorial, 86, 117
- Regressor variables, 13
- REML, 585, 659
- Reparametrizing the Scheffé mixture model, 716
- Replicates, 44
- Residual analysis, 36, 88
- Residuals, 17
- Resolution III design, 164, 165, 188, 195
- Resolution IV design, 165, 197
- Resolution V design, 165, 197
- Response model approach in robust design, 635, 661
- Response optimization, 9, 273, 325
- Response surface designs, 6
- Response surface methodology (RSM), 1, 3, 7
- Response trace plots, 716
- Response variable, 1
- Restricted split-plot designs, 582
- Ridge analysis of a response surface, 289, 290, 291
- Ridge analysis of process variance, 651, 652
- Ridge systems, 282

- Rising ridge system, 283
- Robust design, 12, 619
- Robust design in mixture-process experiments, 778
- Robust parameter design, 619
- Rotatability, 55, 129, 373, 399, 401, 403, 406, 803
- Rotatability and the CCD, 403
- Rotatability of a second-order equiradial design, 803
- RSM designs with qualitative variables, 508, 516
- R*-student, 41
  
- Saturated design, 188, 379, 451, 462
- Scaled prediction variance, 391, 407
- Scheffé mixture model, 698, 716
- Schlaifflian matrices, 797
- Score function, 594
- Screening designs, 375
- Screening experiment, 7, 81, 161
- Screening mixture components, 783
- Second phase of steepest ascent, 244
- Second-order model, 3, 4, 14, 47, 131, 244, 273, 277, 393, 581
- Sequences of fractional factorial designs, 166
- Sequential nature of RSM, 7, 394
- Signal-to-noise ratios, 622, 623, 624, 625
- Simplex coordinate system, 695
- Simplex designs, 381, 383
- Simplex-lattice designs, 696
- Simplex-centroid design, 704
- Simultaneous confidence intervals, 32
- Single replicate of a  $2^k$  factorial, 108
- Slack variable mixture model, 718
- Small composite designs, 458, 484
- Space-filling designs, 567
- Space-filling designs for mixture experiments, 751
- Sphere-packing designs, 569
- Spherical design region, 411
- Split-plot designs, 141, 145, 216, 254, 581
- Split-plot designs for mixture-process experiments, 767
- Squared-error loss methods for multiple response optimization 350
- Standard error of a regression coefficient, 28
- Standard error of effects in a  $2^k$  factorial, 101
- Standard error of the predicted response, 296
  
- Standard order for a  $2^k$  factorial, 85
- Standardized PRESS residuals, 40
- Standardized residuals, 38
- Stationary point, 278, 299
- Stationary ridge system, 282
- Statistical error, 2
- Steepest ascent, 81, 91, 233, 241, 244, 247
- Steepest ascent in a split-plot experiment, 255
- Steepest ascent subject to a linear constraint, 250
- Stochastic computer models, 563
- Studentized residuals, 39, 40
- Subplot factors, 145
- Subplots, 145
- Synergistic blending, 699, 701
  
- Taguchi, G., 12, 619
- Taylor series expansion, 6
- Test for significance of regression, 24
- Testing lack of fit in regression, 43
- Tests on groups of coefficients, 28
- Tests on individual regression coefficients, 27
- Threshold approach, 328
- Transformation, 61, 63
- Transformation of the response, 63, 64, 65
- Two-level fractional factorial designs, 161
  
- Uncertainty in predictions and optimization, 352
- Uniform designs, 569
- Unscaled prediction variance, 128, 407
- Upper bound constraints in mixture experiments, 743
- U*-pseudocomponents, 747
  
- Variance dispersion graphs, 478
- Variance dispersion graphs for cuboidal regions, 485
- Variance optimal design, 376, 377
- Variance stabilizing link, 606
- V*-optimal design, 470, 472
  
- Wald inference on model coefficients in a GLM, 595
- Whole plot, 145
- Whole plot factors, 145
- Words in a defining relation, 178
  
- X** matrix, 16

## WILEY SERIES IN PROBABILITY AND STATISTICS

ESTABLISHED BY WALTER A. SHEWHART AND SAMUEL S. WILKS

Editors: *David J. Balding, Noel A. C. Cressie, Garrett M. Fitzmaurice, Geof H. Givens, Harvey Goldstein, Geert Molenberghs, David W. Scott, Adrian F. M. Smith, Ruey S. Tsay, Sanford Weisberg*  
Editors Emeriti: *J. Stuart Hunter, Iain M. Johnstone, Joseph B. Kadane, Jozef L. Teugels*

The *Wiley Series in Probability and Statistics* is well established and authoritative. It covers many topics of current research interest in both pure and applied statistics and probability theory. Written by leading statisticians and institutions, the titles span both state-of-the-art developments in the field and classical methods.

Reflecting the wide range of current research in statistics, the series encompasses applied, methodological and theoretical statistics, ranging from applications and new techniques made possible by advances in computerized practice to rigorous treatment of theoretical approaches.

This series provides essential and invaluable reading for all statisticians, whether in academia, industry, government, or research.

- † ABRAHAM and LEDOLTER · Statistical Methods for Forecasting  
AGRESTI · Analysis of Ordinal Categorical Data, *Second Edition*  
AGRESTI · An Introduction to Categorical Data Analysis, *Second Edition*  
AGRESTI · Categorical Data Analysis, *Third Edition*  
AGRESTI · *Foundations of Linear and Generalized Linear Models*  
ALSTON, MENGERSEN and PETTITT (editors) · Case Studies in Bayesian Statistical Modelling and Analysis  
ALTMAN, GILL, and McDONALD · Numerical Issues in Statistical Computing for the Social Scientist  
AMARATUNGA and CABRERA · Exploration and Analysis of DNA Microarray and Protein Array Data  
AMARATUNGA, CABRERA, and SHKEDY · Exploration and Analysis of DNA Microarray and Other High-Dimensional Data, *Second Edition*  
ANDÉL · Mathematics of Chance  
ANDERSON · An Introduction to Multivariate Statistical Analysis, *Third Edition*  
\* ANDERSON · The Statistical Analysis of Time Series  
ANDERSON, AUQUIER, HAUCK, OAKES, VANDAELE, and WEISBERG · Statistical Methods for Comparative Studies  
ANDERSON and LOYNES · The Teaching of Practical Statistics  
ARMITAGE and DAVID (editors) · Advances in Biometry  
ARNOLD, BALAKRISHNAN, and NAGARAJA · Records  
\* ARTHANARI and DODGE · Mathematical Programming in Statistics  
AUGUSTIN, COOLEN, DE COOMAN and TROFFAES (editors) · Introduction to Imprecise Probabilities  
\* BAILEY · The Elements of Stochastic Processes with Applications to the Natural Sciences  
BAJORSKI · Statistics for Imaging, Optics, and Photonics  
BALAKRISHNAN and KOUTRAS · Runs and Scans with Applications  
BALAKRISHNAN and NG · Precedence-Type Tests and Applications  
BARNETT · Comparative Statistical Inference, *Third Edition*  
BARNETT · Environmental Statistics  
BARNETT and LEWIS · Outliers in Statistical Data, *Third Edition*

\*Now available in a lower priced paperback edition in the Wiley Classics Library.

†Now available in a lower priced paperback edition in the Wiley-Interscience Paperback Series.

- BARTHOLOMEW, KNOTT, and MOUSTAKI · Latent Variable Models and Factor Analysis: A Unified Approach, *Third Edition*
- BARTOSZYNSKI and NIEWIADOMSKA-BUGAJ · Probability and Statistical Inference, *Second Edition*
- BASILEVSKY · Statistical Factor Analysis and Related Methods: Theory and Applications
- BATES and WATTS · Nonlinear Regression Analysis and Its Applications
- BECHHOFFER, SANTNER, and GOLDSMAN · Design and Analysis of Experiments for Statistical Selection, Screening, and Multiple Comparisons
- BEH and LOMBARDO · Correspondence Analysis: Theory, Practice and New Strategies
- BEIRLANT, GOEGEBEUR, SEGERS, TEUGELS, and DE WAAL · Statistics of Extremes: Theory and Applications
- BELSLEY · Conditioning Diagnostics: Collinearity and Weak Data in Regression
- † BELSLEY, KUH, and WELSCH · Regression Diagnostics: Identifying Influential Data and Sources of Collinearity
- BENDAT and PIERSOL · Random Data: Analysis and Measurement Procedures, *Fourth Edition*
- BERNARDO and SMITH · Bayesian Theory
- BHAT and MILLER · Elements of Applied Stochastic Processes, *Third Edition*
- BHATTACHARYA and WAYMIRE · Stochastic Processes with Applications
- BIEMER, GROVES, LYBERG, MATHIOWETZ, and SUDMAN · Measurement Errors in Surveys
- BILLINGSLEY · Convergence of Probability Measures, *Second Edition*
- BILLINGSLEY · Probability and Measure, *Anniversary Edition*
- BIRKES and DODGE · Alternative Methods of Regression
- BISGAARD and KULAHCI · Time Series Analysis and Forecasting by Example
- BISWAS, DATTA, FINE, and SEGAL · Statistical Advances in the Biomedical Sciences: Clinical Trials, Epidemiology, Survival Analysis, and Bioinformatics
- BLISCHKE and MURTHY (editors) · Case Studies in Reliability and Maintenance
- BLISCHKE and MURTHY · Reliability: Modeling, Prediction, and Optimization
- BLOOMFIELD · Fourier Analysis of Time Series: An Introduction, *Second Edition*
- BOLLEN · Structural Equations with Latent Variables
- BOLLEN and CURRAN · Latent Curve Models: A Structural Equation Perspective
- BONNINI, CORAIN, MAROZZI and SALMASO · Nonparametric Hypothesis Testing: Rank and Permutation Methods with Applications in R
- BOROVKOV · Ergodicity and Stability of Stochastic Processes
- BOSQ and BLANKE · Inference and Prediction in Large Dimensions
- BOULEAU · Numerical Methods for Stochastic Processes
- \* BOX and TIAO · Bayesian Inference in Statistical Analysis
- BOX · Improving Almost Anything, *Revised Edition*
- \* BOX and DRAPER · Evolutionary Operation: A Statistical Method for Process Improvement
- BOX and DRAPER · Response Surfaces, Mixtures, and Ridge Analyses, *Second Edition*
- BOX, HUNTER, and HUNTER · Statistics for Experimenters: Design, Innovation, and Discovery, *Second Edition*
- BOX, JENKINS, REINSEL, and LJUNG · Time Series Analysis: Forecasting and Control, *Fifth Edition*
- BOX, LUCENO, and PANIAGUA-QUINONES · Statistical Control by Monitoring and Adjustment, *Second Edition*
- \* BROWN and HOLLANDER · Statistics: A Biomedical Introduction
- CAIROLI and DALANG · Sequential Stochastic Optimization

\*Now available in a lower priced paperback edition in the Wiley Classics Library.

†Now available in a lower priced paperback edition in the Wiley-Interscience Paperback Series.

- CASTILLO, HADI, BALAKRISHNAN, and SARABIA · Extreme Value and Related Models with Applications in Engineering and Science
- CHAN · Time Series: Applications to Finance with R and S-Plus<sup>®</sup>, *Second Edition*
- CHARALAMBIDES · Combinatorial Methods in Discrete Distributions
- CHATTERJEE and HADI · Regression Analysis by Example, *Fourth Edition*
- CHATTERJEE and HADI · Sensitivity Analysis in Linear Regression
- CHEN · The Fitness of Information: Quantitative Assessments of Critical Evidence
- CHERNICK · Bootstrap Methods: A Guide for Practitioners and Researchers, *Second Edition*
- CHERNICK and FRIIS · Introductory Biostatistics for the Health Sciences
- CHILÈS and DELFINER · Geostatistics: Modeling Spatial Uncertainty, *Second Edition*
- CHIU, STOYAN, KENDALL and MECKE · Stochastic Geometry and Its Applications, *Third Edition*
- CHOW and LIU · Design and Analysis of Clinical Trials: Concepts and Methodologies, *Third Edition*
- CLARKE · Linear Models: The Theory and Application of Analysis of Variance
- CLARKE and DISNEY · Probability and Random Processes: A First Course with Applications, *Second Edition*
- \* COCHRAN and COX · Experimental Designs, *Second Edition*
- COLLINS and LANZA · Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences
- CONGDON · Applied Bayesian Modelling, *Second Edition*
- CONGDON · Bayesian Models for Categorical Data
- CONGDON · Bayesian Statistical Modelling, *Second Edition*
- CONOVER · Practical Nonparametric Statistics, *Third Edition*
- COOK · Regression Graphics
- COOK and WEISBERG · An Introduction to Regression Graphics
- COOK and WEISBERG · Applied Regression Including Computing and Graphics
- CORNELL · A Primer on Experiments with Mixtures
- CORNELL · Experiments with Mixtures, Designs, Models, and the Analysis of Mixture Data, *Third Edition*
- COX · A Handbook of Introductory Statistical Methods
- CRESSIE · Statistics for Spatial Data, *Revised Edition*
- CRESSIE and WIKLE · Statistics for Spatio-Temporal Data
- CSÖRGŐ and HORVÁTH · Limit Theorems in Change Point Analysis
- DAGPUNAR · Simulation and Monte Carlo: With Applications in Finance and MCMC
- DANIEL · Applications of Statistics to Industrial Experimentation
- DANIEL · Biostatistics: A Foundation for Analysis in the Health Sciences, *Eighth Edition*
- \* DANIEL · Fitting Equations to Data: Computer Analysis of Multifactor Data, *Second Edition*
- DASU and JOHNSON · Exploratory Data Mining and Data Cleaning
- DAVID and NAGARAJA · Order Statistics, *Third Edition*
- DAVINO, FURNO and VISTOCCO · Quantile Regression: Theory and Applications
- \* DEGROOT, FIENBERG, and KADANE · Statistics and the Law
- DEL CASTILLO · Statistical Process Adjustment for Quality Control
- DEMARIS · Regression with Social Data: Modeling Continuous and Limited Response Variables
- DEMIDENKO · Mixed Models: Theory and Applications with R, *Second Edition*

\*Now available in a lower priced paperback edition in the Wiley Classics Library.

†Now available in a lower priced paperback edition in the Wiley-Interscience Paperback Series.

- DENISON, HOLMES, MALLICK, and SMITH · Bayesian Methods for Nonlinear Classification and Regression
- DETTE and STUDDEN · The Theory of Canonical Moments with Applications in Statistics, Probability, and Analysis
- DEY and MUKERJEE · Fractional Factorial Plans
- DILLON and GOLDSTEIN · Multivariate Analysis: Methods and Applications
- \* DODGE and ROMIG · Sampling Inspection Tables, *Second Edition*
- \* DOOB · Stochastic Processes
- DOWDY, WEARDEN, and CHILKO · Statistics for Research, *Third Edition*
- DRAPER and SMITH · Applied Regression Analysis, *Third Edition*
- DRYDEN and MARDIA · Statistical Shape Analysis
- DUDEWICZ and MISHRA · Modern Mathematical Statistics
- DUNN and CLARK · Basic Statistics: A Primer for the Biomedical Sciences, *Fourth Edition*
- DUPUIS and ELLIS · A Weak Convergence Approach to the Theory of Large Deviations
- EDLER and KITSOS · Recent Advances in Quantitative Methods in Cancer and Human Health Risk Assessment
- \* ELANDT-JOHNSON and JOHNSON · Survival Models and Data Analysis
- ENDERS · Applied Econometric Time Series, *Third Edition*
- † ETHIER and KURTZ · Markov Processes: Characterization and Convergence
- EVANS, HASTINGS, and PEACOCK · Statistical Distributions, *Third Edition*
- EVERITT, LANDAU, LEESE, and STAHL · Cluster Analysis, *Fifth Edition*
- FEDERER and KING · Variations on Split Plot and Split Block Experiment Designs
- FELLER · An Introduction to Probability Theory and Its Applications, Volume I, *Third Edition*, Revised; Volume II, *Second Edition*
- FITZMAURICE, LAIRD, and WARE · Applied Longitudinal Analysis, *Second Edition*
- \* FLEISS · The Design and Analysis of Clinical Experiments
- FLEISS · Statistical Methods for Rates and Proportions, *Third Edition*
- † FLEMING and HARRINGTON · Counting Processes and Survival Analysis
- FUJIKOSHI, ULYANOV, and SHIMIZU · Multivariate Statistics: High-Dimensional and Large-Sample Approximations
- FULLER · Introduction to Statistical Time Series, *Second Edition*
- † FULLER · Measurement Error Models
- GALLANT · Nonlinear Statistical Models
- GEISSER · Modes of Parametric Statistical Inference
- GELMAN and MENG · Applied Bayesian Modeling and Causal Inference from ncomplete-Data Perspectives
- GEWEKE · Contemporary Bayesian Econometrics and Statistics
- GHOSH, MUKHOPADHYAY, and SEN · Sequential Estimation
- GIESBRECHT and GUMPERTZ · Planning, Construction, and Statistical Analysis of Comparative Experiments
- GIFI · Nonlinear Multivariate Analysis
- GIVENS and HOETING · Computational Statistics
- GLASSERMAN and YAO · Monotone Structure in Discrete-Event Systems
- GNANADESIKAN · Methods for Statistical Data Analysis of Multivariate Observations, *Second Edition*
- GOLDSTEIN · Multilevel Statistical Models, *Fourth Edition*
- GOLDSTEIN and LEWIS · Assessment: Problems, Development, and Statistical Issues
- GOLDSTEIN and WOOFF · Bayes Linear Statistics

\*Now available in a lower priced paperback edition in the Wiley Classics Library.

†Now available in a lower priced paperback edition in the Wiley-Interscience Paperback Series.

- GRAHAM · Markov Chains: Analytic and Monte Carlo Computations  
 GREENWOOD and NIKULIN · A Guide to Chi-Squared Testing  
 GROSS, SHORTLE, THOMPSON, and HARRIS · Fundamentals of Queueing Theory, *Fourth Edition*  
 GROSS, SHORTLE, THOMPSON, and HARRIS · Solutions Manual to Accompany Fundamentals of Queueing Theory, *Fourth Edition*  
 \* HAHN and SHAPIRO · Statistical Models in Engineering  
 HAHN and MEEKER · Statistical Intervals: A Guide for Practitioners  
 HALD · A History of Probability and Statistics and their Applications Before 1750  
 † HAMPEL · Robust Statistics: The Approach Based on Influence Functions  
 HARTUNG, KNAPP, and SINHA · Statistical Meta-Analysis with Applications  
 HEIBERGER · Computation for the Analysis of Designed Experiments  
 HEDAYAT and SINHA · Design and Inference in Finite Population Sampling  
 HEDEKER and GIBBONS · Longitudinal Data Analysis  
 HELLER · MACSYMA for Statisticians  
 HERITIER, CANTONI, COPT, and VICTORIA-FESER · Robust Methods in Biostatistics  
 HINKELMANN and KEMPTHORNE · Design and Analysis of Experiments, Volume 1: Introduction to Experimental Design, *Second Edition*  
 HINKELMANN and KEMPTHORNE · Design and Analysis of Experiments, Volume 2: Advanced Experimental Design  
 HINKELMANN (editor) · Design and Analysis of Experiments, Volume 3: Special Designs and Applications  
 HOAGLIN, MOSTELLER, and TUKEY · Fundamentals of Exploratory Analysis of Variance  
 \* HOAGLIN, MOSTELLER, and TUKEY · Exploring Data Tables, Trends and Shapes  
 \* HOAGLIN, MOSTELLER, and TUKEY · Understanding Robust and Exploratory Data Analysis  
 HOCHBERG and TAMHANE · Multiple Comparison Procedures  
 HOCKING · Methods and Applications of Linear Models: Regression and the Analysis of Variance, *Third Edition*  
 HOEL · Introduction to Mathematical Statistics, *Fifth Edition*  
 HOGG and KLUGMAN · Loss Distributions  
 HOLLANDER, WOLFE, and CHICKEN · Nonparametric Statistical Methods, *Third Edition*  
 HOSMER and LEMESHOW · Applied Logistic Regression, *Second Edition*  
 HOSMER, LEMESHOW, and MAY · Applied Survival Analysis: Regression Modeling of Time-to-Event Data, *Second Edition*  
 HUBER · Data Analysis: What Can Be Learned From the Past 50 Years  
 HUBER · Robust Statistics  
 † HUBER and RONCHETTI · Robust Statistics, *Second Edition*  
 HUBERTY · Applied Discriminant Analysis, *Second Edition*  
 HUBERTY and OLEJNIK · Applied MANOVA and Discriminant Analysis, *Second Edition*  
 HUITEMA · The Analysis of Covariance and Alternatives: Statistical Methods for Experiments, Quasi-Experiments, and Single-Case Studies, *Second Edition*  
 HUNT and KENNEDY · Financial Derivatives in Theory and Practice, *Revised Edition*  
 HURD and MIAMEE · Periodically Correlated Random Sequences: Spectral Theory and Practice  
 HUSKOVA, BERAN, and DUPAC · Collected Works of Jaroslav Hajek— with Commentary  
 HUZURBAZAR · Flowgraph Models for Multistate Time-to-Event Data

\*Now available in a lower priced paperback edition in the Wiley Classics Library.

†Now available in a lower priced paperback edition in the Wiley-Interscience Paperback Series.



- JACKMAN · Bayesian Analysis for the Social Sciences
- † JACKSON · A User's Guide to Principle Components
- JOHN · Statistical Methods in Engineering and Quality Assurance
- JOHNSON · Multivariate Statistical Simulation
- JOHNSON and BALAKRISHNAN · Advances in the Theory and Practice of Statistics: A Volume in Honor of Samuel Kotz
- JOHNSON, KEMP, and KOTZ · Univariate Discrete Distributions, *Third Edition*
- JOHNSON and KOTZ (editors) · Leading Personalities in Statistical Sciences: From the Seventeenth Century to the Present
- JOHNSON, KOTZ, and BALAKRISHNAN · Continuous Univariate Distributions, Volume 1, *Second Edition*
- JOHNSON, KOTZ, and BALAKRISHNAN · Continuous Univariate Distributions, Volume 2, *Second Edition*
- JOHNSON, KOTZ, and BALAKRISHNAN · Discrete Multivariate Distributions
- JUDGE, GRIFFITHS, HILL, LÜTKEPOHL, and LEE · The Theory and Practice of Econometrics, *Second Edition*
- JUREK and MASON · Operator-Limit Distributions in Probability Theory
- KADANE · Bayesian Methods and Ethics in a Clinical Trial Design
- KADANE AND SCHUM · A Probabilistic Analysis of the Sacco and Vanzetti Evidence
- KALBFLEISCH and PRENTICE · The Statistical Analysis of Failure Time Data, *Second Edition*
- KARIYA and KURATA · Generalized Least Squares
- KASS and VOS · Geometrical Foundations of Asymptotic Inference
- † KAUFMAN and ROUSSEEUW · Finding Groups in Data: An Introduction to Cluster Analysis
- KEDEM and FOKIANOS · Regression Models for Time Series Analysis
- KENDALL, BARDEN, CARNE, and LE · Shape and Shape Theory
- KHURI · Advanced Calculus with Applications in Statistics, *Second Edition*
- KHURI, MATHEW, and SINHA · Statistical Tests for Mixed Linear Models
- \* KISH · Statistical Design for Research
- KLEIBER and KOTZ · Statistical Size Distributions in Economics and Actuarial Sciences
- KLEMELÄ · Smoothing of Multivariate Data: Density Estimation and Visualization
- KLUGMAN, PANJER, and WILLMOT · Loss Models: From Data to Decisions, *Third Edition*
- KLUGMAN, PANJER, and WILLMOT · Loss Models: Further Topics
- KLUGMAN, PANJER, and WILLMOT · Solutions Manual to Accompany Loss Models: From Data to Decisions, *Third Edition*
- KOSKI and NOBLE · Bayesian Networks: An Introduction
- KOTZ, BALAKRISHNAN, and JOHNSON · Continuous Multivariate Distributions, Volume 1, *Second Edition*
- KOTZ and JOHNSON (editors) · Encyclopedia of Statistical Sciences: Volumes 1 to 9 with Index
- KOTZ and JOHNSON (editors) · Encyclopedia of Statistical Sciences: Supplement Volume
- KOTZ, READ, and BANKS (editors) · Encyclopedia of Statistical Sciences: Update Volume 1
- KOTZ, READ, and BANKS (editors) · Encyclopedia of Statistical Sciences: Update Volume 2
- KOWALSKI and TU · Modern Applied U-Statistics
- KRISHNAMOORTHY and MATHEW · Statistical Tolerance Regions: Theory, Applications, and Computation

\*Now available in a lower priced paperback edition in the Wiley Classics Library.

†Now available in a lower priced paperback edition in the Wiley-Interscience Paperback Series.

- KROESE, TAIMRE, and BOTEV · Handbook of Monte Carlo Methods
- KROONENBERG · Applied Multiway Data Analysis
- KULINSKAYA, MORGENTHALER, and STAUDTE · Meta Analysis: A Guide to Calibrating and Combining Statistical Evidence
- KULKARNI and HARMAN · An Elementary Introduction to Statistical Learning Theory
- KUROWICKA and COOKE · Uncertainty Analysis with High Dimensional Dependence Modelling
- KVAM and VIDAKOVIC · Nonparametric Statistics with Applications to Science and Engineering
- LACHIN · Biostatistical Methods: The Assessment of Relative Risks, *Second Edition*
- LAD · Operational Subjective Statistical Methods: A Mathematical, Philosophical, and Historical Introduction
- LAMPERTI · Probability: A Survey of the Mathematical Theory, *Second Edition*
- LAWLESS · Statistical Models and Methods for Lifetime Data, *Second Edition*
- LAWSON · Statistical Methods in Spatial Epidemiology, *Second Edition*
- LE · Applied Categorical Data Analysis, *Second Edition*
- LE · Applied Survival Analysis
- LEE · Structural Equation Modeling: A Bayesian Approach
- LEE and WANG · Statistical Methods for Survival Data Analysis, *Fourth Edition*
- LEPAGE and BILLARD · Exploring the Limits of Bootstrap
- LESSLER and KALSBECK · Nonsampling Errors in Surveys
- LEYLAND and GOLDSTEIN (editors) · Multilevel Modelling of Health Statistics
- LIAO · Statistical Group Comparison
- LIN · Introductory Stochastic Analysis for Finance and Insurance
- LINDLEY · Understanding Uncertainty, *Revised Edition*
- LITTLE and RUBIN · Statistical Analysis with Missing Data, *Second Edition*
- LLOYD · The Statistical Analysis of Categorical Data
- LOWEN and TEICH · Fractal-Based Point Processes
- MAGNUS and NEUDECKER · Matrix Differential Calculus with Applications in Statistics and Econometrics, *Revised Edition*
- MALLER and ZHOU · Survival Analysis with Long Term Survivors
- MARCHETTE · Random Graphs for Statistical Pattern Recognition
- MARDIA and JUPP · Directional Statistics
- MARKOVICH · Nonparametric Analysis of Univariate Heavy-Tailed Data: Research and Practice
- MARONNA, MARTIN and YOHAJ · Robust Statistics: Theory and Methods
- MASON, GUNST, and HESS · Statistical Design and Analysis of Experiments with Applications to Engineering and Science, *Second Edition*
- McCULLOCH, SEARLE, and NEUHAUS · Generalized, Linear, and Mixed Models, *Second Edition*
- McFADDEN · Management of Data in Clinical Trials, *Second Edition*
- \* McLACHLAN · Discriminant Analysis and Statistical Pattern Recognition
- McLACHLAN, DO, and AMBROISE · Analyzing Microarray Gene Expression Data
- McLACHLAN and KRISHNAN · The EM Algorithm and Extensions, *Second Edition*
- McLACHLAN and PEEL · Finite Mixture Models
- McNEIL · Epidemiological Research Methods
- MEEKER and ESCOBAR · Statistical Methods for Reliability Data
- MEERSCHAERT and SCHEFFLER · Limit Distributions for Sums of Independent Random Vectors: Heavy Tails in Theory and Practice
- MENGERSEN, ROBERT, and TITTERINGTON · Mixtures: Estimation and Applications

\*Now available in a lower priced paperback edition in the Wiley Classics Library.

†Now available in a lower priced paperback edition in the Wiley-Interscience Paperback Series.

- MICKEY, DUNN, and CLARK · Applied Statistics: Analysis of Variance and Regression, *Third Edition*
- \* MILLER · Survival Analysis, *Second Edition*
- MONTGOMERY, JENNINGS, and KULAHCI · Introduction to Time Series Analysis and Forecasting, *Second Edition*
- MONTGOMERY, PECK, and VINING · Introduction to Linear Regression Analysis, *Fifth Edition*
- MORGENTHALER and TUKEY · Configural Polysampling: A Route to Practical Robustness
- MUIRHEAD · Aspects of Multivariate Statistical Theory
- MULLER and STOYAN · Comparison Methods for Stochastic Models and Risks
- MURTHY, XIE, and JIANG · Weibull Models
- MYERS, MONTGOMERY, and ANDERSON-COOK · Response Surface Methodology: Process and Product Optimization Using Designed Experiments, *Fourth Edition*
- MYERS, MONTGOMERY, VINING, and ROBINSON · Generalized Linear Models. With Applications in Engineering and the Sciences, *Second Edition*
- NATVIG · Multistate Systems Reliability Theory With Applications
- † NELSON · Accelerated Testing, Statistical Models, Test Plans, and Data Analyses
- † NELSON · Applied Life Data Analysis
- NEWMAN · Biostatistical Methods in Epidemiology
- NG, TAIN, and TANG · Dirichlet Theory: Theory, Methods and Applications
- OKABE, BOOTS, SUGIHARA, and CHIU · Spatial Tessellations: Concepts and Applications of Voronoi Diagrams, *Second Edition*
- OLIVER and SMITH · Influence Diagrams, Belief Nets and Decision Analysis
- PALMA · Time Series Analysis
- PALTA · Quantitative Methods in Population Health: Extensions of Ordinary Regressions
- PANJER · Operational Risk: Modeling and Analytics
- PANKRATZ · Forecasting with Dynamic Regression Models
- PANKRATZ · Forecasting with Univariate Box-Jenkins Models: Concepts and Cases
- PARDOUX · Markov Processes and Applications: Algorithms, Networks, Genome and Finance
- PARMIGIANI and INOUE · Decision Theory: Principles and Approaches
- \* PARZEN · Modern Probability Theory and Its Applications
- PEÑA, TIAO, and TSAY · A Course in Time Series Analysis
- PESARIN and SALMASO · Permutation Tests for Complex Data: Applications and Software
- PIANTADOSI · Clinical Trials: A Methodologic Perspective, *Second Edition*
- POURAHMADI · Foundations of Time Series Analysis and Prediction Theory
- POURAHMADI · High-Dimensional Covariance Estimation
- POWELL · Approximate Dynamic Programming: Solving the Curses of Dimensionality, *Second Edition*
- POWELL and RYZHOV · Optimal Learning
- PRESS · Subjective and Objective Bayesian Statistics, *Second Edition*
- PRESS and TANUR · The Subjectivity of Scientists and the Bayesian Approach
- PURI, VILAPLANA, and WERTZ · New Perspectives in Theoretical and Applied Statistics
- † PUTERMAN · Markov Decision Processes: Discrete Stochastic Dynamic Programming
- QIU · Image Processing and Jump Regression Analysis
- \* RAO · Linear Statistical Inference and Its Applications, *Second Edition*

\*Now available in a lower priced paperback edition in the Wiley Classics Library.

†Now available in a lower priced paperback edition in the Wiley-Interscience Paperback Series.

- RAO · Statistical Inference for Fractional Diffusion Processes
- RAUSAND and HØYLAND · System Reliability Theory: Models, Statistical Methods, and Applications, *Second Edition*
- RAYNER, THAS, and BEST · Smooth Tests of Goodness of Fit: Using R, *Second Edition*
- RENCHEr and SCHAALJE · Linear Models in Statistics, *Second Edition*
- RENCHEr and CHRISTENSEN · Methods of Multivariate Analysis, *Third Edition*
- RENCHEr · Multivariate Statistical Inference with Applications
- RIGDON and BASU · Statistical Methods for the Reliability of Repairable Systems
- \* RIPLEY · Spatial Statistics
- \* RIPLEY · Stochastic Simulation
- ROHATGI and SALEH · An Introduction to Probability and Statistics, *Third Edition*
- ROLSKI, SCHMIDLI, SCHMIDT, and TEUGELS · Stochastic Processes for Insurance and Finance
- ROSENBERGER and LACHIN · Randomization in Clinical Trials: Theory and Practice
- ROSSI, ALLENBY, and McCULLOCH · Bayesian Statistics and Marketing
- † ROUSSEEUW and LEROY · Robust Regression and Outlier Detection
- ROYSTON and SAUERBREI · Multivariate Model Building: A Pragmatic Approach to Regression Analysis Based on Fractional Polynomials for Modeling Continuous Variables
- \* RUBIN · Multiple Imputation for Nonresponse in Surveys
- RUBINSTEIN and KROESE · Simulation and the Monte Carlo Method, *Second Edition*
- RUBINSTEIN and MELAMED · Modern Simulation and Modeling
- RUBINSTEIN, RIDDER, and VAISMAN · Fast Sequential Monte Carlo Methods for Counting and Optimization
- RYAN · Modern Engineering Statistics
- RYAN · Modern Experimental Design
- RYAN · Modern Regression Methods, *Second Edition*
- RYAN · Sample Size Determination and Power
- RYAN · Statistical Methods for Quality Improvement, *Third Edition*
- SALEH · Theory of Preliminary Test and Stein-Type Estimation with Applications
- SALTELLI, CHAN, and SCOTT (editors) · Sensitivity Analysis
- SCHERER · Batch Effects and Noise in Microarray Experiments: Sources and Solutions
- \* SCHEFFE · The Analysis of Variance
- SCHIMEK · Smoothing and Regression: Approaches, Computation, and Application
- SCHOTT · Matrix Analysis for Statistics, *Second Edition*
- SCHOUTENS · Levy Processes in Finance: Pricing Financial Derivatives
- SCOTT · Multivariate Density Estimation
- SCOTT · Multivariate Density Estimation: Theory, Practice, and Visualization
- \* SEARLE · Linear Models
- † SEARLE · Linear Models for Unbalanced Data
- † SEARLE · Matrix Algebra Useful for Statistics
- † SEARLE, CASELLA, and McCULLOCH · Variance Components
- SEARLE and WILLETT · Matrix Algebra for Applied Economics
- SEBER · A Matrix Handbook For Statisticians
- † SEBER · Multivariate Observations
- SEBER and LEE · Linear Regression Analysis, *Second Edition*
- † SEBER and WILD · Nonlinear Regression

\*Now available in a lower priced paperback edition in the Wiley Classics Library.

†Now available in a lower priced paperback edition in the Wiley-Interscience Paperback Series.

- SENNOTT · Stochastic Dynamic Programming and the Control of Queuing Systems
- \* SERFLING · Approximation Theorems of Mathematical Statistics
- SHAFFER and VOVK · Probability and Finance: It's Only a Game!
- SHERMAN · Spatial Statistics and Spatio-Temporal Data: Covariance Functions and Directional Properties
- SILVAPULLE and SEN · Constrained Statistical Inference: Inequality, Order, and Shape Restrictions
- SINGPURWALLA · Reliability and Risk: A Bayesian Perspective
- SMALL and MCLEISH · Hilbert Space Methods in Probability and Statistical Inference
- SRIVASTAVA · Methods of Multivariate Statistics
- STAPLETON · Linear Statistical Models, *Second Edition*
- STAPLETON · Models for Probability and Statistical Inference: Theory and Applications
- STAUDTE and SHEATHER · Robust Estimation and Testing
- STOYAN · Counterexamples in Probability, *Second Edition*
- STOYAN and STOYAN · Fractals, Random Shapes and Point Fields: Methods of Geometrical Statistics
- STREET and BURGESS · The Construction of Optimal Stated Choice Experiments: Theory and Methods
- STYAN · The Collected Papers of T. W. Anderson: 1943–1985
- SUTTON, ABRAMS, JONES, SHELDON, and SONG · Methods for Meta-Analysis in Medical Research
- TAKEZAWA · Introduction to Nonparametric Regression
- TAMHANE · Statistical Analysis of Designed Experiments: Theory and Applications
- TANAKA · Time Series Analysis: Nonstationary and Noninvertible Distribution Theory
- THOMPSON · Empirical Model Building: Data, Models, and Reality, *Second Edition*
- THOMPSON · Sampling, *Third Edition*
- THOMPSON · Simulation: A Modeler's Approach
- THOMPSON and SEBER · Adaptive Sampling
- THOMPSON, WILLIAMS, and FINDLAY · Models for Investors in Real World Markets
- TIERNEY · LISP-STAT: An Object-Oriented Environment for Statistical Computing and Dynamic Graphics
- TROFFAES and DE COOMAN · Lower Previsions
- TSAY · Analysis of Financial Time Series, *Third Edition*
- TSAY · An Introduction to Analysis of Financial Data with R
- TSAY · Multivariate Time Series Analysis: With R and Financial Applications
- UPTON and FINGLETON · Spatial Data Analysis by Example, Volume II: Categorical and Directional Data
- † VAN BELLE · Statistical Rules of Thumb, *Second Edition*
- VAN BELLE, FISHER, HEAGERTY, and LUMLEY · Biostatistics: A Methodology for the Health Sciences, *Second Edition*
- VESTRUP · The Theory of Measures and Integration
- VIDAKOVIC · Statistical Modeling by Wavelets
- VIERTL · Statistical Methods for Fuzzy Data
- VINOD and REAGLE · Preparing for the Worst: Incorporating Downside Risk in Stock Market Investments
- WALLER and GOTWAY · Applied Spatial Statistics for Public Health Data

\*Now available in a lower priced paperback edition in the Wiley Classics Library.

†Now available in a lower priced paperback edition in the Wiley–Interscience Paperback Series.

- WEISBERG · Applied Linear Regression, *Fourth Edition*
- WEISBERG · Bias and Causation: Models and Judgment for Valid Comparisons
- WELSH · Aspects of Statistical Inference
- WESTFALL and YOUNG · Resampling-Based Multiple Testing: Examples and Methods for  $p$ -Value Adjustment
- \* WHITTAKER · Graphical Models in Applied Multivariate Statistics
- WINKER · Optimization Heuristics in Economics: Applications of Threshold Accepting
- WOODWORTH · Biostatistics: A Bayesian Introduction
- WOOLSON and CLARKE · Statistical Methods for the Analysis of Biomedical Data, *Second Edition*
- WU and HAMADA · Experiments: Planning, Analysis, and Parameter Design Optimization, *Second Edition*
- WU and ZHANG · Nonparametric Regression Methods for Longitudinal Data Analysis
- YAKIR · Extremes in Random Fields
- YIN · Clinical Trial Design: Bayesian and Frequentist Adaptive Methods
- YOUNG, VALERO-MORA, and FRIENDLY · Visual Statistics: Seeing Data with Dynamic Interactive Graphics
- ZACKS · Examples and Problems in Mathematical Statistics
- ZACKS · Stage-Wise Adaptive Designs
- \* ZELLNER · An Introduction to Bayesian Inference in Econometrics
- ZELTERMAN · Discrete Distributions—Applications in the Health Sciences
- ZHOU, OBUCHOWSKI, and McCLISH · Statistical Methods in Diagnostic Medicine, *Second Edition*

\*Now available in a lower priced paperback edition in the Wiley Classics Library.

†Now available in a lower priced paperback edition in the Wiley–Interscience Paperback Series.

# **WILEY END USER LICENSE AGREEMENT**

Go to [www.wiley.com/go/eula](http://www.wiley.com/go/eula) to access Wiley's ebook  
EULA.