VPM's
DR VN BRIMS, Thane
Programme: MMS (2014-16)
First Semester Examination December 2014

Subject	**Business Statistics (BS 01)**		
Roll No.		**Marks**	**60 Marks**
Total No. of Questions	**7**	**Duration**	**3 Hours**
Total No. of printed pages		**Date**	**08-12-2014**

Note: Q1 is compulsory and solve any FOUR from the remaining SIX questions.

Q1) 20 Marks (Compulsory)

Q.1.a] Multiple choice questions. Write the correct option in the answer sheets.

1. A sample of 50 respondents were asked to give their opinion on their gaming experience pre and post a software update of a popular game application of google play services. The respondents were given a "ten point rating scale" to judge the difference. The proper statistical test for examining the data is

   a. Regression analysis b. ANOVA
   b. t-test for 2 Dependent Samples d. Z-test

2. D mart wants to see if sales (quantitative variable) of a particular breakfast cereal brand differs when it is placed on either of the three shelfs i.e top, middle and bottom (labeled as S1, S2, S3). The researchers have have taken weekly sales of past 5 weeks for the three shelves. Following is the data:

Top Shelf (Sales Unit)	S1	50	60	70	60	55
Middle Shelf (Sales Unit)	S2	80	85	90	75	65
Bottom Shelf (Sales Unit)	S3	40	45	35	40	25

   The type of the test for analysis is:

   a. One way analysis of variance b. Z test c. Multiple Regression d.Chi Square

3. A consultant performed regression analysis for purchase factors for a laptop, where he found r = 0.8. The model is
   a. Weak b. Moderate c. Strong

4. A study of discrimination at work against pregnant women is conducted by the HR students @ DR VN BRIMS. They have surveyed the respondents belonging to junior, middle and senior

management cadre (Qualitative Variable). One of the questions asked was as:

Rank the following instances that you have faced in your organisation (1 signifying highest severity whereas 5 signifying lowest severity of the discrimination) a. Have faced discrimination at work during recruitment process b. Have faced discrimination at work during appraisal process c. Have faced discrimination at work during training and development recommendation process d. Have faced discrimination at work during allocation of key projects to employees

This is best described as an example of _______________ data.

a. Ordinal b. Ratio c. Interval d. Nominal

5. The variable "management cadre – Junior, middle and senior" can be summarised using mean.

a. True b. False

6. For a normally distributed data set ***"About 68.27% of the values lie within 1 standard deviation of the mean. Similarly, about 95.45% of the values lie within 2 standard deviations of the mean. Nearly all (99.73%) of the values lie within 3 standard deviations of the mean."***

The above result is an outcome of which of the followings principles in statistics

a. Chebeyshev's Lemma b. Central Limit Theorem
c. Empirical Rule d. Baye's Theorem

7. The Special Rule of Multiplication is used when the two events are

a. Independent Events b. Mutually Exclusive events

b. Events based on subjective probabilities d. None of the above

8, 9 & 10) ***Note:*** *This question is for 3 marks and No marks will be given if calculations are not shown*

For a sample of 100 stocks traded yesterday on the American Stock Exchange, 25 showed a decline of $1.00, 25 showed no change, and 50 increased by $2.00. Find the weighted mean of the ***price change***

a) 0.65 b) 0.75 c) 50 d) 25

A.1.b)

(i) The following figures relate to production in kilograms of three varieties A, B and C of what sown on 12 plots **(8 Marks)**

A	14	16	18		
B	14	13	15	22	
C	18	16	19	19	20

a) Calculate the grand mean.
b) Estimate between column variance
c) Calculate within column variance
d) Calculate F ration . Is there any significant difference in the production of three varieties if the table value of F is given as 1.13.

(ii) Write a note on application of analysis of variance in business management. **(2 Marks)**

Attempt Any FOUR from the Remaining SIX Questions

Q2) ***Any two from (a) or (b) or (c) ————— (5x2) = 10 Marks***

Illustration: The Air Transport Association of India publishes figures on the busiest airports in the United States. Data is about the number of passengers arriving and departing on various airports. The following frequency distribution has been constructed from these figures for a recent year.

No. Of Passengers (Millions)	No.Of Airports
20-30	8
30-40	7
40-50	1
50-60	0
60-70	3
70-80	1

a) All the three Quartiles and the box plot using the above illustration

b) Standard Deviation and Coefficient of Variation using the above illustration

c) Calculate the coefficient of rank correlation from the following data.

X:	48	33	40	09	16	16	65	24	16	57
Y:	13	13	24	06	15	04	20	09	06	19

Q3) ***Any two from (a) or (b) or (c) ————— (5x2) = 10 Marks***
a) Purchasing Survey asked purchasing professionals what sales traits impressed them most in a sales representative. 78% selected “thoroughness.” 40% percent responded “knowledge of your own product.” and 27% of the purchasing professionals listed both “thoroughness” and “knowledge of your

own product" as sales traits that impressed them most. A purchasing professional is randomly sampled.

i. What is the probability that the professional selected "thoroughness" or "knowledge of your own product"?
ii. What is the probability that the professional selected neither "thoroughness" nor "knowledge of your own product"?

b) In a survey, 70% of all companies are classified as small companies and the rest as large companies. Further, 82% of large companies provide training to employees, but only 18% of small companies provide training. A company is randomly selected. It is known that the company provides training to employees. Using *bayes theorem* what is the probability that this company is

i. a large company
ii. a small company

c) Write a note on application of probability techniques in business management

Q4) ***Any two from (a) or (b) or (c) ————— (5x2) = 10 Marks***

a) Assume baggage is rarely lost by Northwest Airlines. Suppose a random sample of 1,000 flights shows a total of 300 bags were lost. Thus, the arithmetic mean number of lost bags per flight is 0.3 (300/1,000). If the number of lost bags per flight follows a *Poisson distribution* with m= 0.3,

i. What is the probability that no bags will be lost on a particular flight.
ii. What is the probability exactly one bag will be lost on a particular flight?

b) The amount of mustard dispensed from a machine at The Hotdog Emporium is *normally distributed* with a mean of 0.9 ounce and a standard deviation of 0.1 ounce.
i. What is the probability that the mustard dispensed is with in the range of 0.6 and 1.2 ounces
ii. If the machine is used 500 times, *approximately* how many times will it be expected to dispense 1 or more ounces of mustard.

c) Assume that a school district has 10,000 6th graders. In this district, the average weight of a 6th grader is 80 pounds, with a standard deviation of 20 pounds. Suppose you *draw a random sample of 50* students.

i. Can central limit theorem be applied in this case to find the below probability ? If Yes Why ? If No Why?
ii. What is the probability that the average weight of a sampled student will be less than 75 pounds?

Q5) ***Any two from (a) or (b) or (c) ————— (5x2) = 10 Marks***

a) The management claims that the average cash deposit is $200.00, and you've taken a random sample to test that:

192	188	152	211	201	167	177	191	178	209
185	226	192	190	156	224	191	203	186	160

At the 0.05 significance level, test the hypothesis that the average of all cash deposits is less than $200?

b) This question has two sub questions of 3 and 2 marks each
(i) A clothing company produces men's jeans. The jeans are made and sold with either a regular cut or a boot cut. In an effort to estimate the proportion of their men's jeans market in Oklahoma City that prefers boot-cut jeans, the analyst takes a random sample of 212 jeans sales from the company's two Oklahoma City retail outlets. Only 34 of the sales were for boot-cut jeans. Construct a 95% confidence interval to estimate the proportion of the population in Oklahoma City who prefer boot-cut jeans. (3Marks)

(ii) Define "Consumer Surplus" and "Matrix" (2 marks)

c) Write a note on various non probabilistic sampling techniques

Q6) ***Any two from (a) or (b) or (c) ————— (5x2) = 10 Marks***

a) ***<u>In the following cases, (1) State the null and alternative hypothesis, (2) State the test to be used for a particular case. (3) State the decision as to accept or reject Ho (4) State the business Implication</u>***

According to Runzheimer International, a family of four in Manhattan with $60,000 annual income spends more than $22,000 a year on basic goods and services. In contrast, a family of four in San Antonio with the same annual income spends only $15,460 on the same items.

Suppose we want to test to determine whether the variance of money spent per year on the basics by families across the United States is greater than the variance of money spent on the basics by families in Manhattan—that is, whether the amounts spent by families of four in Manhattan are more homogeneous than the amounts spent by such families nationally.

The statistical analysis output is as follows

Calculated value = 8.09 p-Value = 0.0042

Can the variance of values taken from across the United States can be shown to be greater than the variance of values obtained from families in Manhattan . Let = 0.01. Assume the amount spent on the basics is normally distributed in the population.

b) Is the type of beverage ordered with lunch at a restaurant independent of the age of the consumer? A random poll of 309 lunch customers is taken, resulting in the following contingency table of observed values.

	Coffee / Tea	Soft drink	Others (Milk Etc)	Row Total
21-35	26	95	18	139
36-55	41	40	20	101

> 55	24	13	32	69
Colomn Total	91	148	70	309

Using chi-square analysis determine whether the two variables are related at level of significance = 0.01

Given that, $\chi^2_{Cal} = 13.27$

c) Write a note on characteristics of Binomial, Poisson, and Normal Distribution

Q7) ***Any two from (a) or (b) or (c) ————— (5x2) = 10 Marks***

a). Solve the following equations using determinants

$x + y + z - 7 = 0$
$x + 2y + 3z - 16 = 0$
$x + 3y + 4z - 22 = 0$

b). Mr. A purchased equal quantity of sharpener, pencil and rubber aggregating to 15 units. Sharpner, pencil and rubber costs Rs. 14, Rs. 12 and Rs. 7 respectively. Using the matrix multiplication determine the total cost of the amount spent by Mr. A.

c) A hotel has 40 rooms available to be given on rent at the rate of Rs. 1000 per day. However, for each Rs. 100 per day increase in rent two rooms will remain vacant. How many rooms should be rent per day and what rate to maximise the daily income?

============================= End =============================

Sr. No	Concept	Formula
1	**Summation** 7. Summation of n numbers	$\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + \ldots\ldots + x_n$ Simply as, $\sum x = x_1 + x_2 + x_3 + \ldots\ldots + x_n$
2	Arithmetic Mean a. **Ungrouped Data** i) **Population Mean** ii) **Sample Mean** b. **Grouped Data**	$\mu = \frac{\sum x}{N}$ $\acute{x} = \frac{\sum x}{n}$ $\acute{x} = \frac{\sum f * x}{\sum f}$
3	Weighted Mean	$\acute{x}_w = \frac{\sum x * w}{\sum w}$
4	Geometric Mean	G.M = $\sqrt[n]{Product\ of\ the\ n\ values} - 1$
5	Median a. **Ungrouped Data** b. **Grouped Data**	$\left(\frac{n+1}{2}\right)^{th} observation \in the\ data\ set$ $Median = L + \left(\frac{\frac{N}{2} - C.F.}{F}\right)(i)$ L: Lower Limit of the Median class N: Total of all frequencies C.F: Cum. Freq of the class preceding to the median class F: Frequency of the median class i : Class Interval
6	Mode c. **Ungrouped Data** d. **Grouped Data**	Most Frequently Occurring observation in the data set

		f ¿ 0 ¿ $1 - f$ ¿ (¿ $+ (f_1 - f_2)$ ¿¿ $)*(i)$ ¿ $\underline{(f_1 - f_0)}$ ¿ $Mode = L +$ ¿ L: Lower Limit of the Modal class i : Class Interval f_1 : Frequency of the Modal Class f_2 : Frequency of the class after the modal class f_0 : Frequency of the class before the modal class
1	Quartiles	$$Q_M = L + \left(\frac{\frac{M*N}{4} - C.F.}{F} \right)(i)$$ Q_M: M[th] Quartile L: Lower Limit of the Quartile Class i : Class Interval N: Total of all frequencies C.F: Cum. Freq of the class preceding to the quartile class F: Frequency of the quartile class
2	Deciles	$$D_M = L + \left(\frac{\frac{M*N}{10} - C.F.}{F} \right)(i)$$ D_M: M[th] Decile L: Lower Limit of the Decile Class i: Class Interval N: Total of all frequencies C.F: Cum. Freq of the class **<u>preceding</u>** to the decile class F: Frequency of the quartile class
3	Percentiles	$$P_M = L + \left(\frac{\frac{M*N}{100} - C.F.}{F} \right)(i)$$
4	Range	Max - Min
5	Inter-quartile Range	$Q_3 - Q_1$
6	Quartile Deviation	$$Quartile\ Deviation = \frac{Q3 - Q1}{2}$$
7	Coeff of QD	$$Coeff\ of\ QD = \frac{Q3 - Q1}{Q3 + Q1}$$
8	Mean Deviation	$$MD = \frac{\sum \lvert x - \acute{x} \rvert}{N}$$
9	Standard Deviation (S.D.)	

		$\sigma=\sqrt{\frac{\sum f*x^2}{\sum f}-\left(\frac{\sum f*x}{\sum f}\right)^2}$ $\sigma=\sqrt{\frac{\sum f*x^2}{\sum f}-\acute{x}^2}$
10	Variance	σ^2
11	Coefficient of Variation	$CV=\frac{\sigma}{\mu}*100$
12	Coefficient of Skewness b. Pearson's c. Bowley's	$S_k = 3\,(\mu - M) / \sigma$ μ: Arithmetic Mean, σ: Standard Deviation M: Median $S_k = \left(\frac{Q3+Q1-2*Q2}{Q3-Q1}\right)$
13	Combination (selecting r components from n components – *Order of selection <u>is not important</u>*)	${}^n_rC$ read as n - C – r ${}^n_rC=\frac{n!}{r!*(n-r)!}$
14	Permutations (selecting r components from n components – *Order of selection **<u>is important</u>***)	${}^n_rP$ read as n - P – r ${}^n_rP=\frac{n!}{(n-r)!}$
15	Addition Rule 2. General 3. Special (Mutually exclusive)	P(A or B) = P(A) + P(B) – P(A and B) P(A or B) = P(A) + P(B)
16	Multiplication Rule 4. General 5. Special (Independent)	P (A and B) = P (A / B) * P(B) = P(B/A) * P(A) P (A and B) = P (A) * P(B)
17	Bayes Theorem	$\boldsymbol{P(A_1 \mid B)} = \frac{P(A_1)P(B/A_1)}{P(A_1)P(B/A_1)+P(A_2)P(B/A_2)}$
18	Expected value	μ = $E(x)=\sum x_i*P(x_i)$
19	Variance	σ² = $V(x)=\sum$ ¿ x ¿ ¿i−¿ ¿ ¿ ¿
20	Binomial Distribution	$P(x=r)={}^n_rC*p^r*q^{n-r}$ Where n: no. of trials p: probability of success

		q: 1-p
21	Poisson Distribution	$$P(x=r)=\frac{m^r * e^{-m}}{x!}$$ n: no. of observations r: no. of expected successes p : Probability of the success m = $n*p$
22	Normal Distribution	$$P(X<x)=\frac{1}{\sqrt{2\pi\sigma}} * e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma^2}\right)}$$
23	Rules for solving Normal Distribution Problem: ------------------------------------ P(a), P(a_1), and P(a_2) to be directly noted from the table ------------------------------------	a. P (Z < a) = 0.5 + P(a) b. P(Z > a) = 0.5 – P(a) c. P(a_1 < Z < a_2) = P(a_2) – P(a_1) d. P(- a1 < Z < a_2) = P(a_2) + P(a_1)
24	Sample Mean $\acute{x} \rightarrow Normal\,Distribution$	$$Z=\frac{\acute{x}-\mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}$$
25	Sample proportion $\hat{p} \rightarrow Normal\,Distribution$	$$Z=\frac{\hat{p}-p}{\left(\sqrt{\frac{p*q}{n}}\right)}$$ Where $\hat{p}$: sample proportion P: Population proportion q: 1 – p n: sample size
26	Confidence Interval for Population Mean (μ)	$$\left(\acute{x}-\frac{\sigma}{\sqrt{n}}*Z,\ \acute{x}+\frac{\sigma}{\sqrt{n}}*Z\right)$$
27	Confidence Interval for Population Proportion (p)	$$\left(p-\sqrt{\frac{p*q}{n}}*Z,\ p+\sqrt{\frac{p*q}{n}}*Z\right)$$
28	Sample Size ❖ Mean ❖ Proportion	$$n=\left(\frac{Z*\sigma}{E}\right)^2$$ $$n=p*q\left(\frac{Z}{E}\right)^2$$
29	Testing of Hypothesis: Sample Mean $\acute{x} \rightarrow Normal\,Distribution$	$$Z=\frac{\acute{x}-\mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}$$
30	Testing of Hypothesis: Sample proportion $\hat{p} \rightarrow Normal\,Distribution$	$$Z=\frac{\hat{p}-p}{\left(\sqrt{\frac{p*q}{n}}\right)}$$

		Where $\hat{p}$: sample proportion P: Population proportion q: 1 – p n: sample size
31	Pearson's Correlation Coefficient	$r=\frac{\left(\frac{\sum x*y}{n}-\acute{x}\acute{y}\right)}{\sigma_x\sigma_y}$
32	Spearman's Rank Correlation Coefficient	$\rho=1-\frac{6\sum d^2}{n(n^2-1)}$
33	Regression equation of • X on Y • Y on X	$(x-\acute{x})=(\frac{r*\sigma_x}{\sigma_y})(y-\acute{y})$ $(y-\acute{y})=(\frac{r*\sigma_y}{\sigma_x})(x-\acute{x})$
34	Chi Square	$\chi^2_{cal}=\frac{\sum(O-E)^2}{E}$