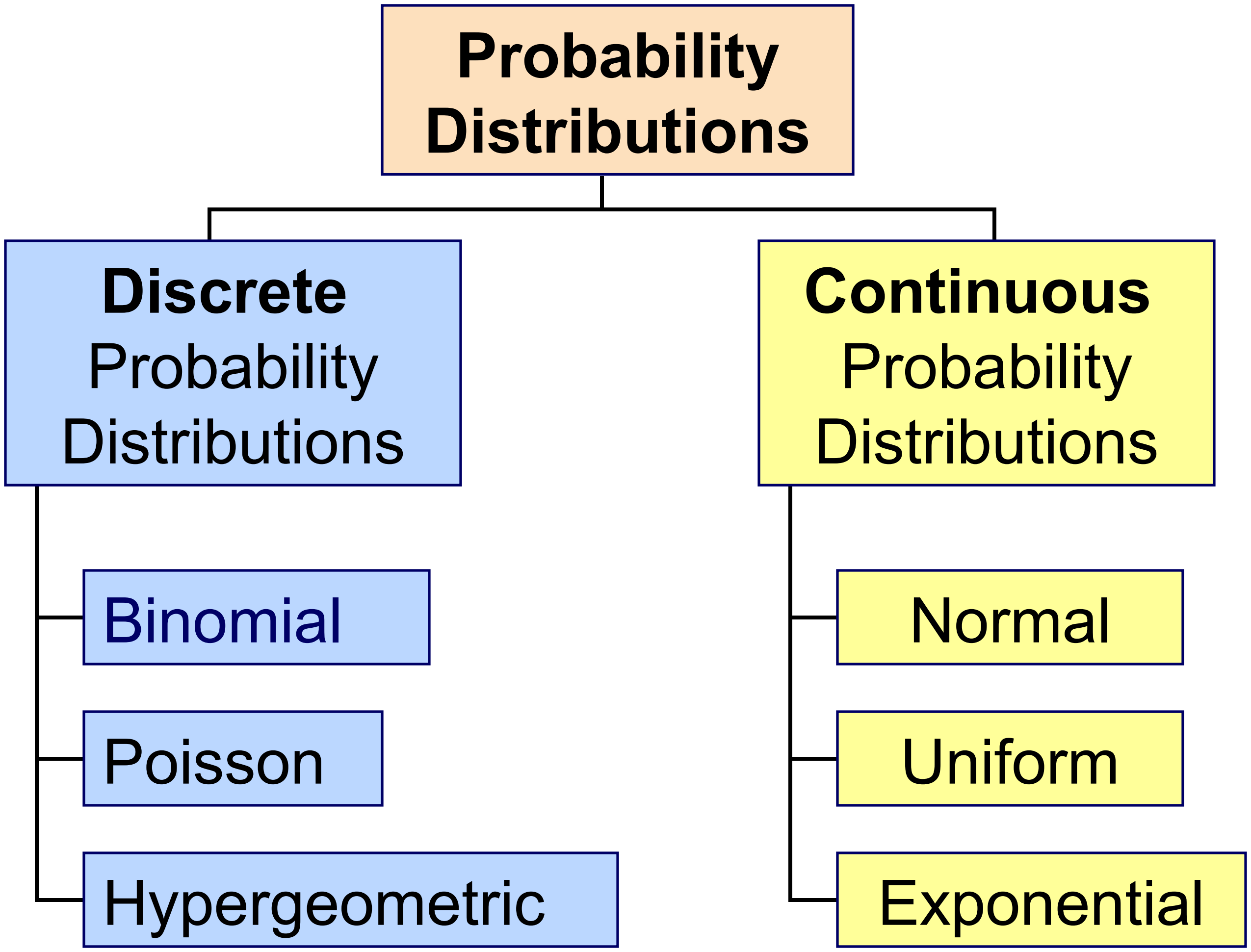
DISCRETE AND CONTINUOUS PROBABILITY DISTRIBUTIONS

PROBABILITY DISTRIBUTION

- It describes how the outcomes of an experiment are expected to vary

Probability Distributions



DISCRETE PROBABILITY DISTRIBUTIONS

- A discrete random variable is a variable that can assume only a countable number of values

Many possible outcomes:

- number of complaints per day
- number of TV's in a household
- number of rings before the phone is answered

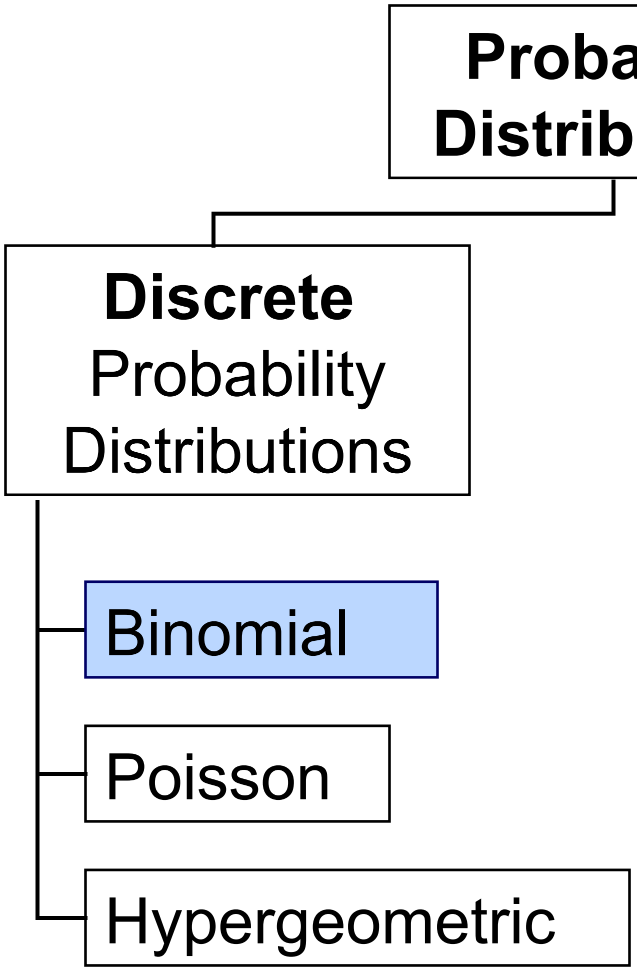
Only two possible outcomes:

- gender: male or female
- defective: yes or no
- spreads peanut butter first vs. spreads jelly first

CONTINUOUS PROBABILITY DISTRIBUTIONS

- A continuous random variable is a variable that can assume any value on a continuum (can assume an uncountable number of values)
  - thickness of an item
  - time required to complete a task
  - temperature of a solution
  - height, in inches
- These can potentially take on any value, depending only on the ability to measure accurately.

The Binomial Distribution



THE BINOMIAL DISTRIBUTION

- Characteristics of the Binomial Distribution:
  - A trial has only two possible outcomes – “success” or “failure”
  - There is a fixed number, n, of identical trials
  - The trials of the experiment are independent of each other
  - The probability of a success, p, remains constant from trial to trial
  - If p represents the probability of a success, then $(1-p) = q$ is the probability of a failure

BINOMIAL DISTRIBUTION SETTINGS

- A manufacturing plant labels items as either defective or acceptable
- A firm bidding for a contract will either get the contract or not
- A marketing research firm receives survey responses of "yes I will buy" or "no I will not"
- New job applicants either accept the offer or reject it

COUNTING RULE FOR COMBINATIONS

- A combination is an outcome of an experiment where x objects are selected from a group of n objects

$$C_x^n = \frac{n!}{x!(n-x)!}$$

where:

$n! = n(n - 1)(n - 2) \ldots (2)(1)$

$x! = x(x - 1)(x - 2) \ldots (2)(1)$

$0! = 1$ (by definition)

BINOMIAL DISTRIBUTION FORMULA

$$P(x) = \frac{n!}{x!(n-x)!} p^{x} q^{n-x}$$

P(x) = probability of **x** successes in **n** trials, with probability of success **p** on each trial

x = number of 'successes' in sample, ($x = 0, 1, 2, ..., n$)

p = probability of "success" per trial

q = probability of "failure" = (1 – p)

n = number of trials (sample size)

Example: Flip a coin four times, let x = # heads:

n = 4

p = 0.5

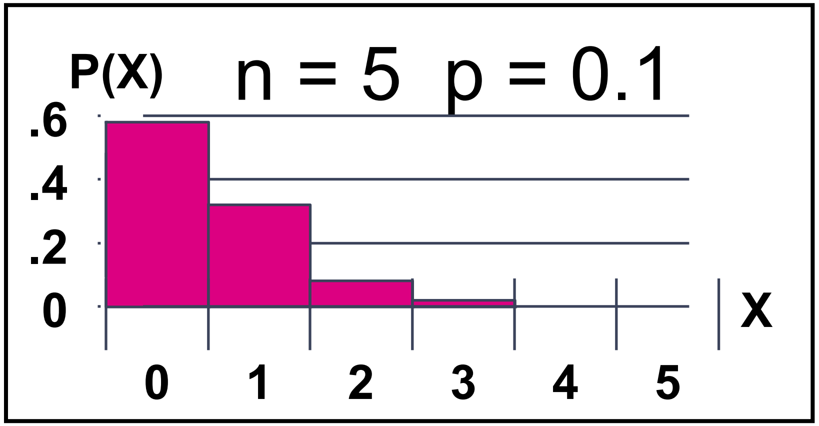
q = (1 - .5) = .5

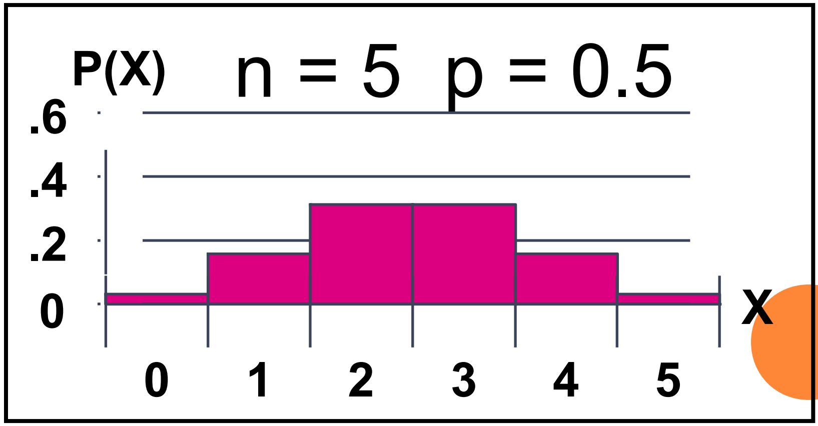
x = 0, 1, 2, 3, 4

BINOMIAL DISTRIBUTION

- The shape of the binomial distribution depends on the values of p and n

Here, n = 5 and p = .1





Here, n = 5 and p = .5

BINOMIAL DISTRIBUTION CHARACTERISTICS

- Mean

$$\mu = E(x) = np$$

Variance and Standard Deviation

$$\sigma^2 = npq$$

$$\sigma = \sqrt{npq}$$

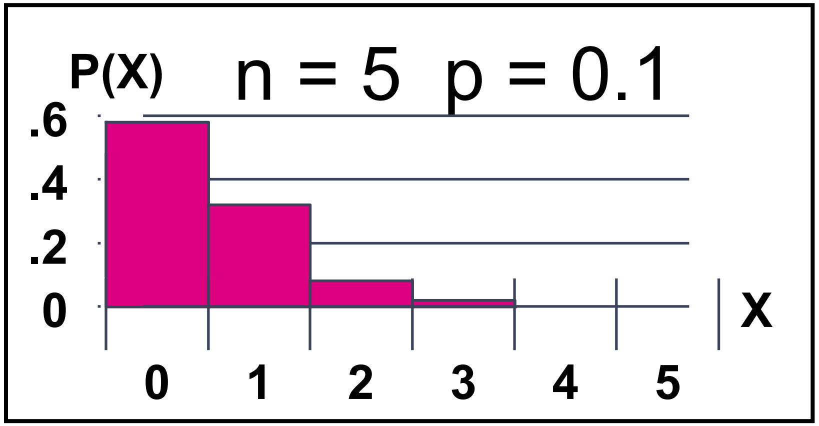
Where n = sample size
p = probability of success
$q = (1 - p)$ = probability of failure

BINOMIAL CHARACTERISTICS

Examples

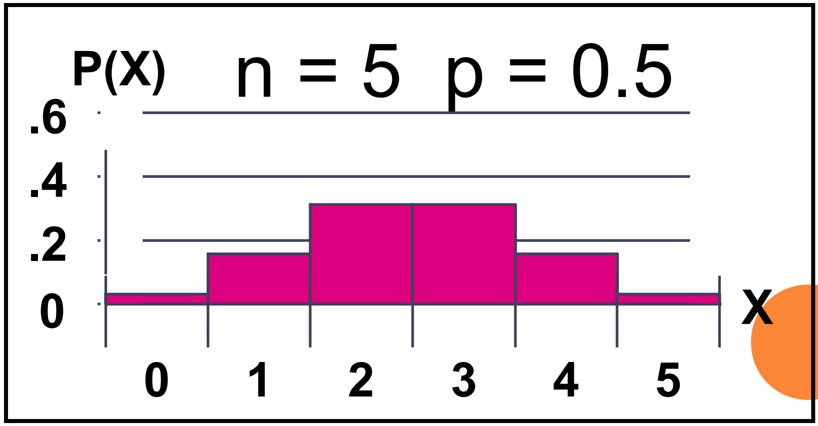
$$\mu = np = (5)(.1) = 0.5$$

$$\sigma = \sqrt{npq} = \sqrt{(5)(.1)(1-.1)}$$
$$= 0.6708$$



$$\mu = np = (5)(.5) = 2.5$$

$$\sigma = \sqrt{npq} = \sqrt{(5)(.5)(1-.5)}$$
$$= 1.118$$



USING BINOMIAL TABLES

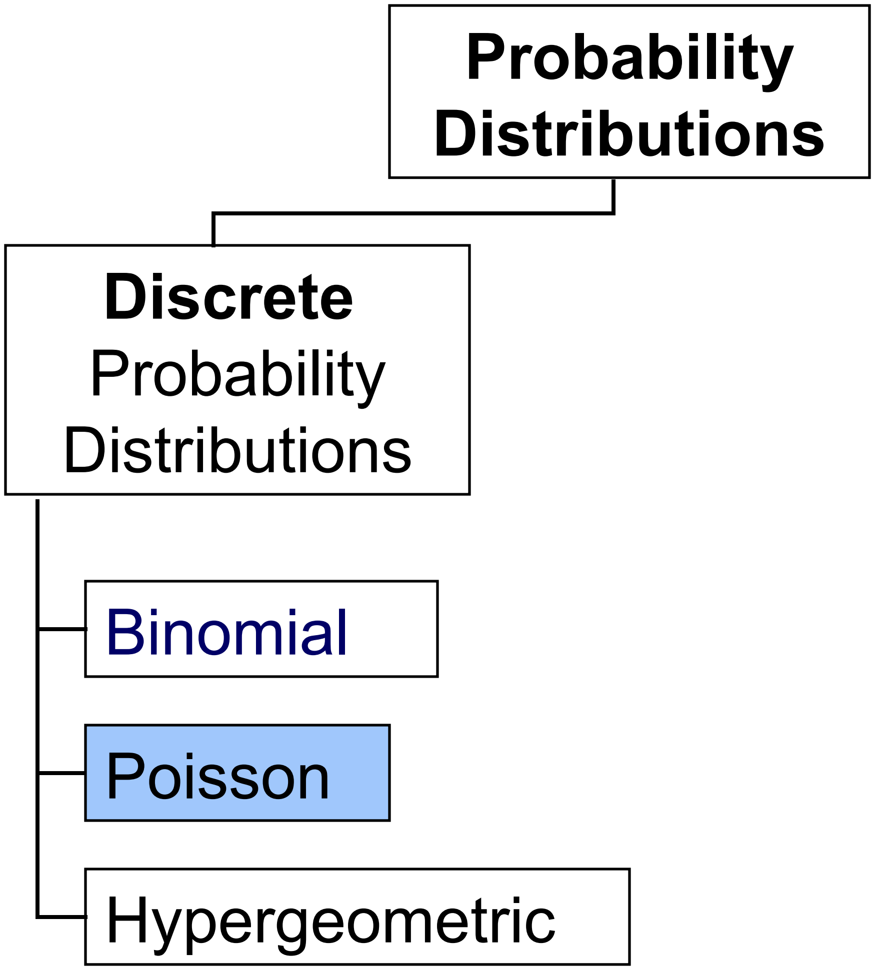
n = 10									
x	p=.15	p=.20	p=.25	p=.30	**p=.35**	p=.40	p=.45	p=.50	
0	0.1969	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010	10
1	0.3474	0.2684	0.1877	0.1211	0.0725	0.0403	0.0207	0.0098	9
2	0.2759	0.3020	0.2816	0.2335	0.1757	0.1209	0.0763	0.0439	8
3	0.1298	0.2013	0.2503	0.2668	**0.2522**	0.2150	0.1665	0.1172	7
4	0.0401	0.0881	0.1460	0.2001	0.2377	0.2508	0.2384	0.2051	6
5	0.0085	0.0264	0.0584	0.1029	0.1536	0.2007	0.2340	0.2461	5
6	0.0012	0.0055	0.0162	0.0368	0.0689	0.1115	0.1596	0.2051	4
7	0.0001	0.0008	0.0031	0.0090	0.0212	0.0425	0.0746	0.1172	3
8	0.0000	0.0001	**0.0004**	0.0014	0.0043	0.0106	0.0229	0.0439	**2**
9	0.0000	0.0000	0.0000	0.0001	0.0005	0.0016	0.0042	0.0098	1
10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010	0
	p=.85	p=.80	**p=.75**	p=.70	p=.65	p=.60	p=.55	p=.50	**x**

Examples:

n = 10, p = .35, x = 3: $P(x = 3|n = 10, p = .35) = .2522$

n = 10, p = .75, x = 2: $P(x = 2|n = 10, p = .75) = .0004$

The Poisson Distribution



THE POISSON DISTRIBUTION

- Characteristics of the Poisson Distribution:
  - The outcomes of interest are rare relative to the possible outcomes
  - The average number of outcomes of interest per time or space interval is **“λ”**
  - The number of outcomes of interest are random, and the occurrence of one outcome does not influence the chances of another outcome of interest
  - The probability of that an outcome of interest occurs in a given segment is the same for all segments

POISSON DISTRIBUTION FORMULA

$$P(x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}$$

where:

t = size of the segment of interest

x = number of successes in segment of interest

λ = expected number of successes in a segment of unit size

e = base of the natural logarithm system (2.71828...)

POISSON DISTRIBUTION CHARACTERISTICS

- Mean

$$\mu = \lambda t$$

Variance and Standard Deviation

$$\sigma^2 = \lambda t$$

$$\sigma = \sqrt{\lambda t}$$

where λ = number of successes in a segment of unit size

t = the size of the segment of interest

USING POISSON TABLES

X	λt								
	0.10	0.20	0.30	0.40	**0.50**	0.60	0.70	0.80	0.90
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.0905	0.1637	0.2222	0.2681	0.3033	0.3293	0.3476	0.3595	0.3659
2	0.0045	0.0164	0.0333	0.0536	**0.0758**	0.0988	0.1217	0.1438	0.1647
3	0.0002	0.0011	0.0033	0.0072	0.0126	0.0198	0.0284	0.0383	0.0494
4	0.0000	0.0001	0.0003	0.0007	0.0016	0.0030	0.0050	0.0077	0.0111
5	0.0000	0.0000	0.0000	0.0001	0.0002	0.0004	0.0007	0.0012	0.0020
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0003
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Example: Find $P(x = 2)$ if $\lambda = .05$ and $t = 10$

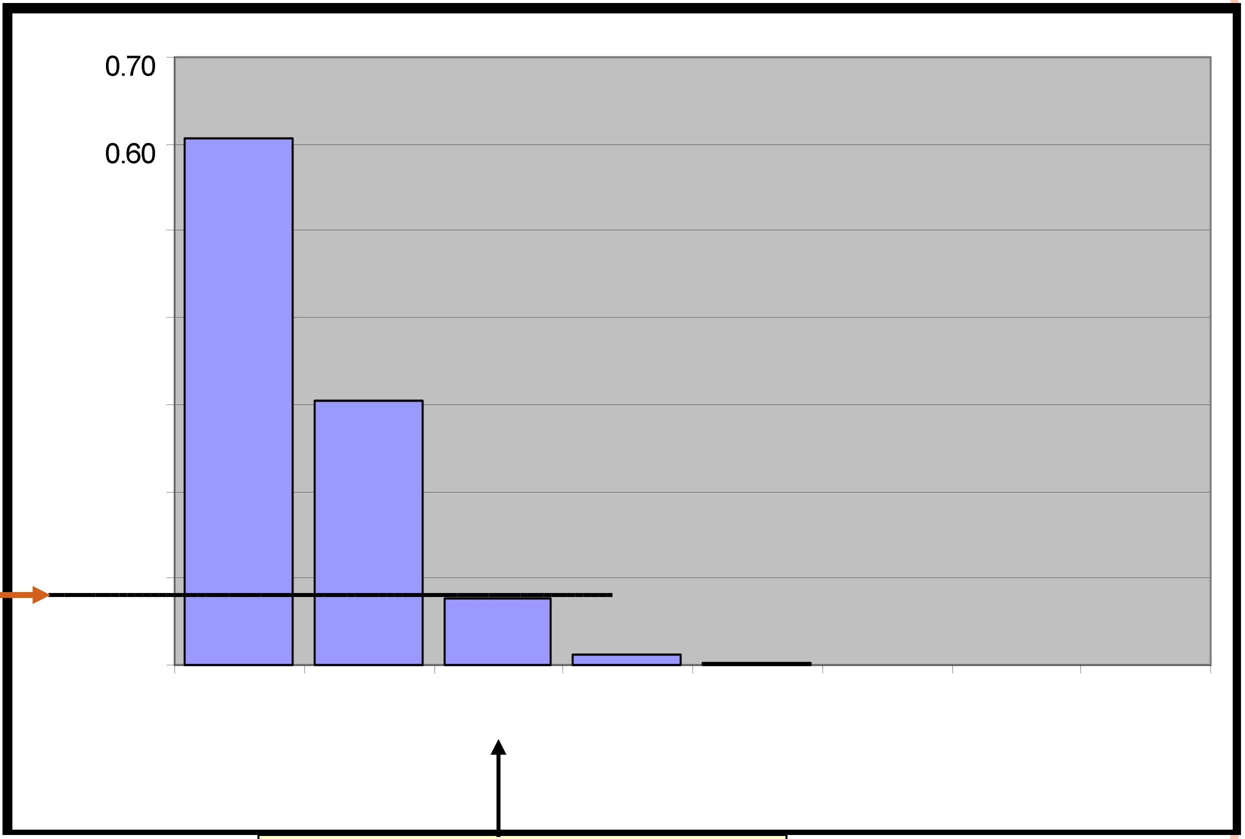
$$P(x=2) = \frac{(\lambda t)^x e^{-\lambda t}}{x!} = \frac{(0.50)^2 e^{-0.50}}{2!} = .0758$$

GRAPH OF POISSON PROBABILITIES

Graphically:

λ = .05 and t = 100

X	**λt = 0.50**
0	0.6065
1	0.3033
2	0.0758
3	0.0126
4	0.0016
5	0.0002
6	0.0000
7	0.0000

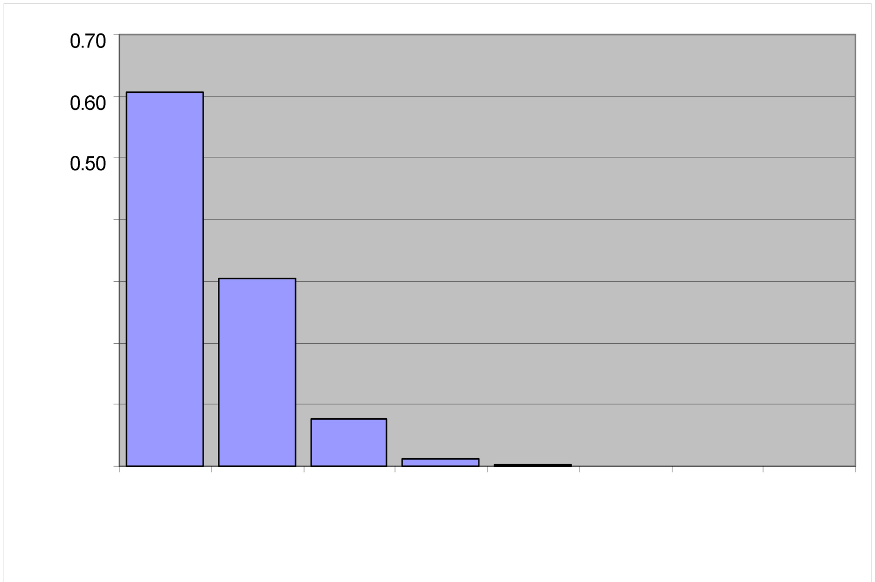


P(x = 2) = .0758

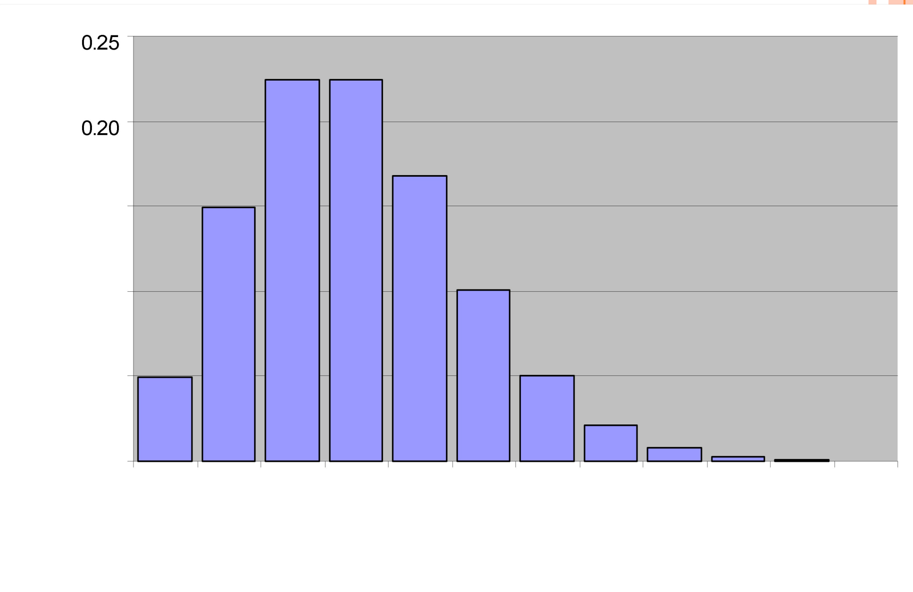
POISSON DISTRIBUTION SHAPE

- The shape of the Poisson Distribution depends on the parameters λ and t:

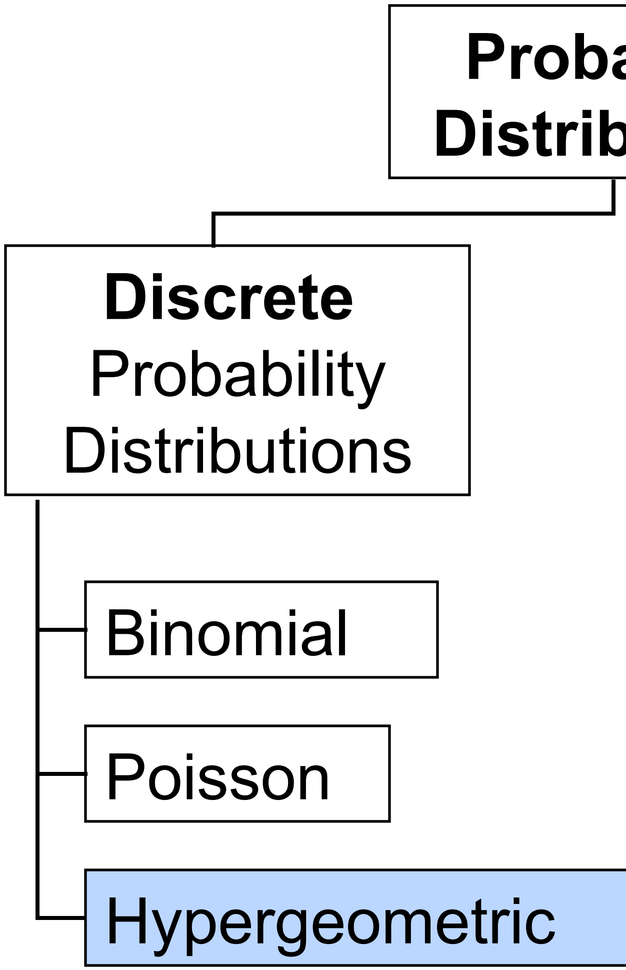
$\lambda t = 0.50$



$\lambda t = 3.0$



The Hypergeometric Distribution



THE HYPERGEOMETRIC DISTRIBUTION

- "n" trials in a sample taken from a finite population of size N
- Sample taken without replacement
- Trials are dependent
- Concerned with finding the probability of "x" successes in the sample where there are "X" successes in the population

HYPERGEOMETRIC DISTRIBUTION FORMULA

(Two possible outcomes per trial)

$$P(x) = \frac{C_{n-x}^{N-X} \cdot C_{x}^{X}}{C_{n}^{N}}$$

Where

N = Population size

X = number of successes in the population

n = sample size

x = number of successes in the sample

n – x = number of failures in the sample

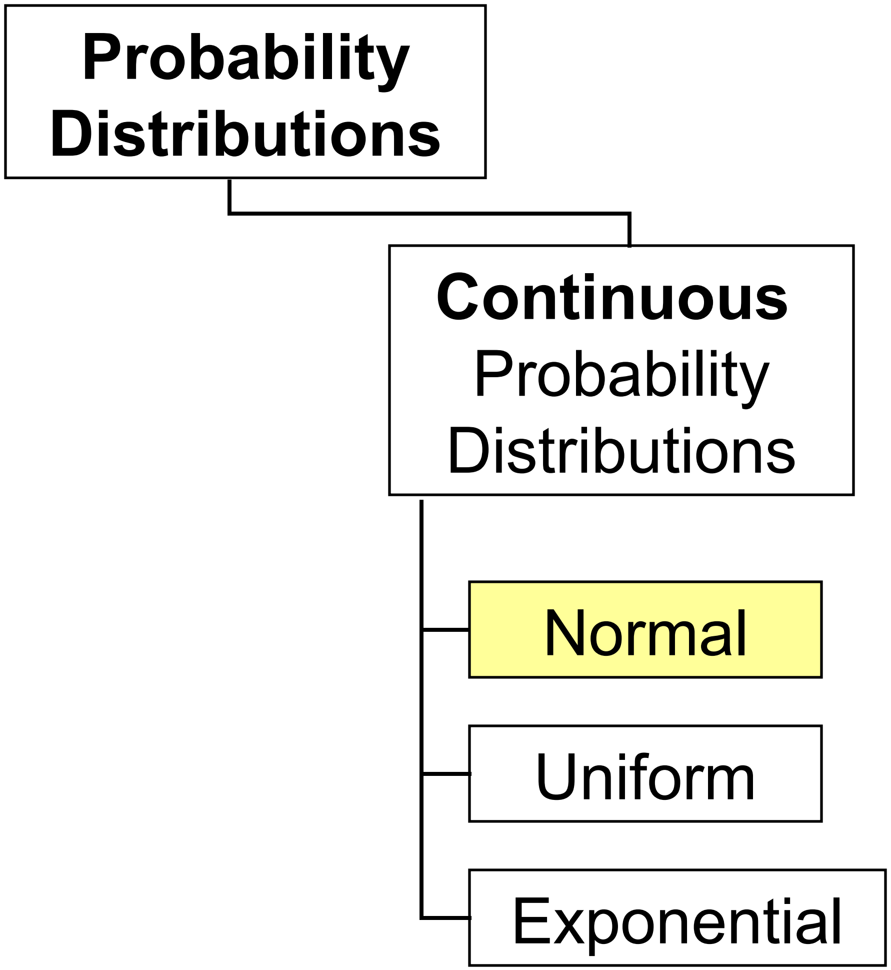
HYPERGEOMETRIC DISTRIBUTION FORMULA

- Example: 3 Light bulbs were selected from 10. Of the 10 there were 4 defective. What is the probability that 2 of the 3 selected are defective?

N = 10 n = 3
X = 4 x = 2

$$P(x=2) = \frac{C_{n-x}^{N-X} \, C_{x}^{X}}{C_{n}^{N}} = \frac{C_{1}^{6} \, C_{2}^{4}}{C_{3}^{10}} = \frac{(6)(6)}{120} = 0.3$$

The Normal Distribution



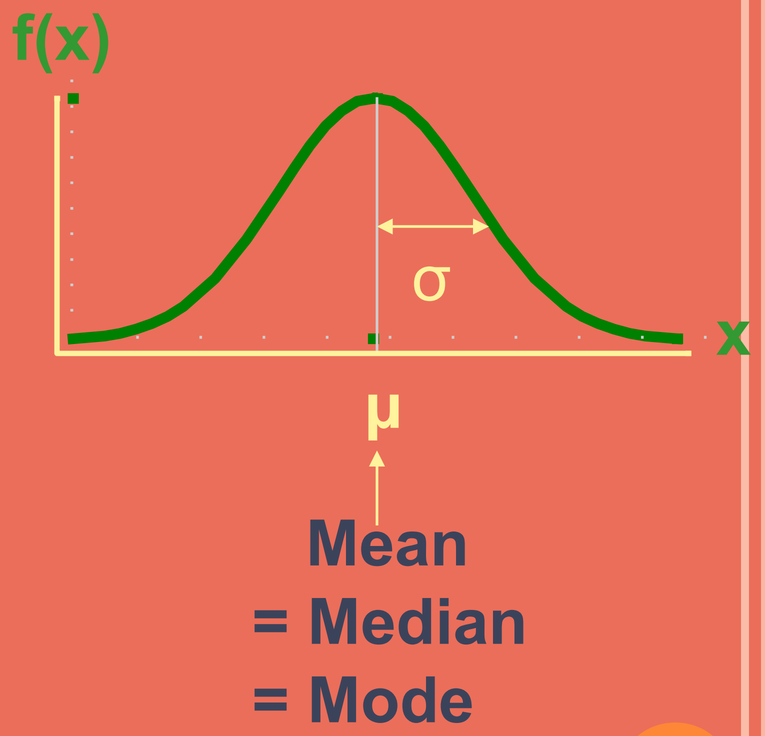
THE NORMAL DISTRIBUTION

- 'Bell Shaped'
- Symmetrical
- Mean, Median and Mode are Equal

Location is determined by the mean, μ

Spread is determined by the standard deviation, σ

The random variable has an infinite theoretical range: $+\infty$ to $-\infty$



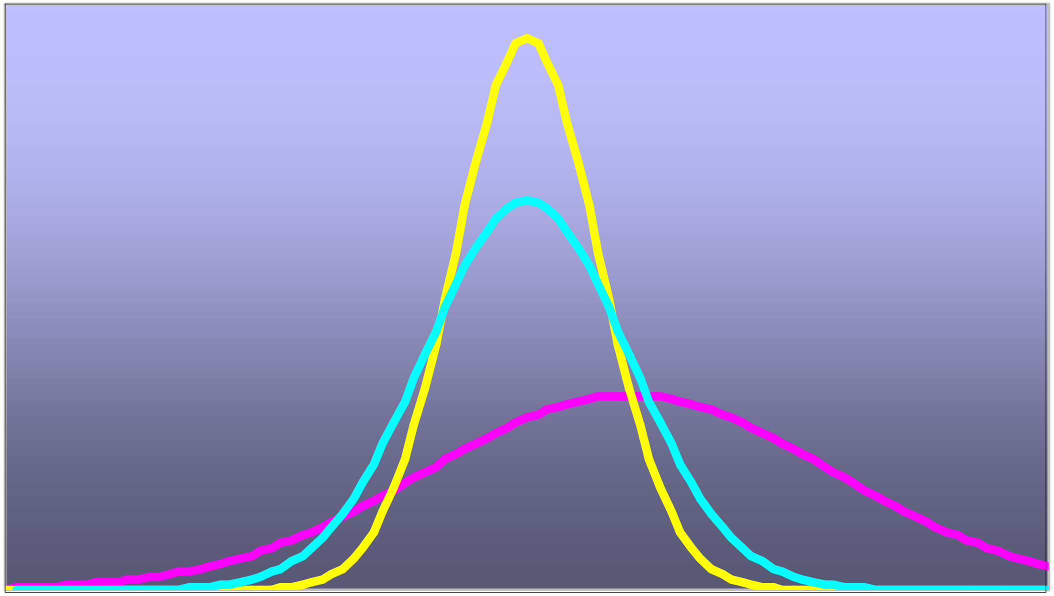
The general formula for the probability density function of the normal distribution is

$$f(x) = \frac{e^{-(x-\mu)^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}}$$

Where μ is mean

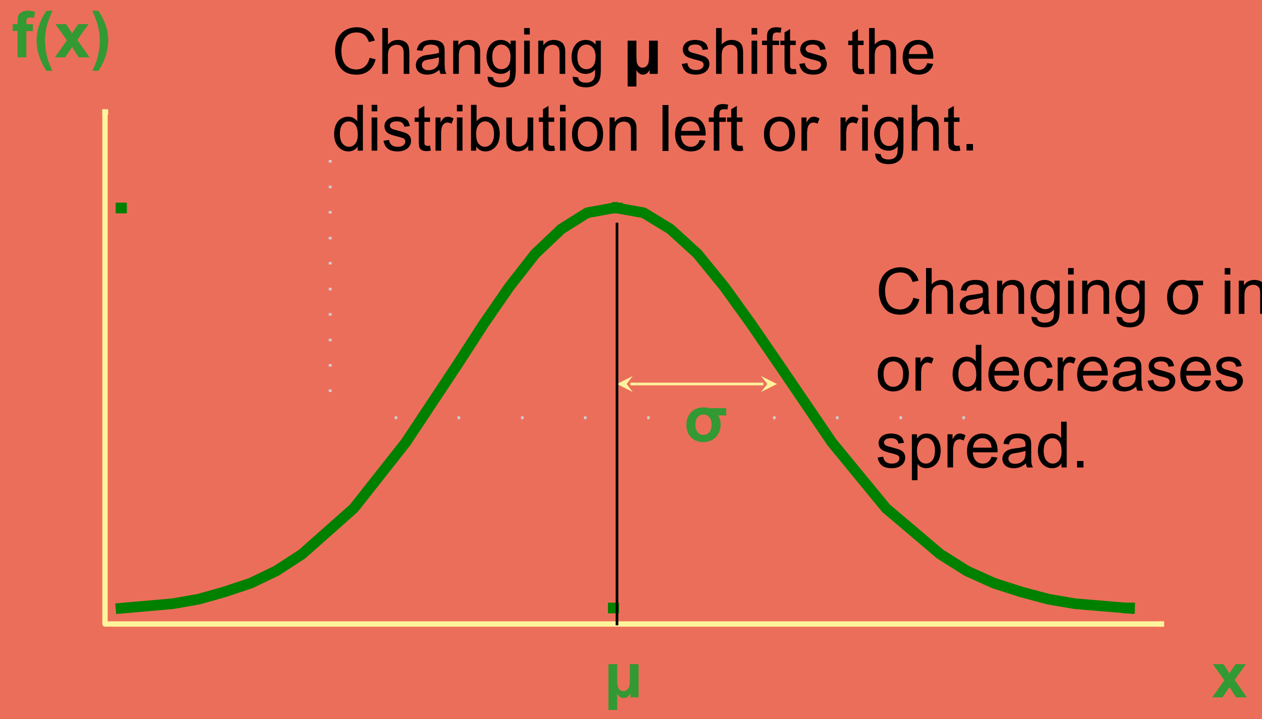
and σ is standard deviation

Many Normal Distributions



By varying the parameters μ and σ, we obtain different normal distributions

THE NORMAL DISTRIBUTION SHAPE

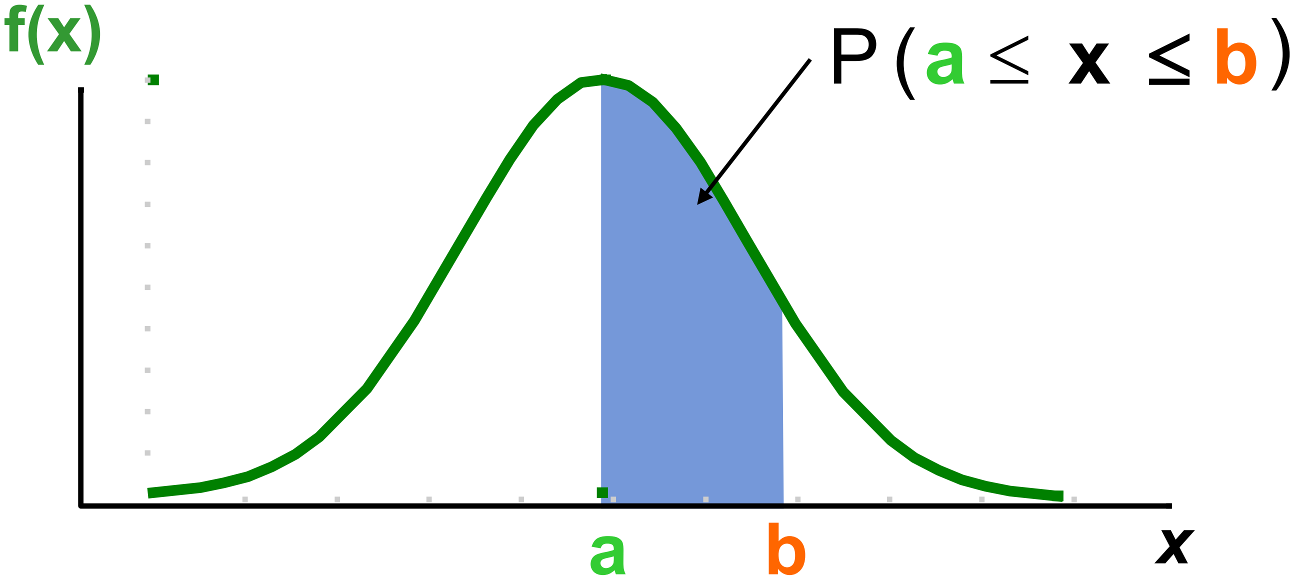


Changing **μ** shifts the distribution left or right.

Changing σ increases or decreases the spread.

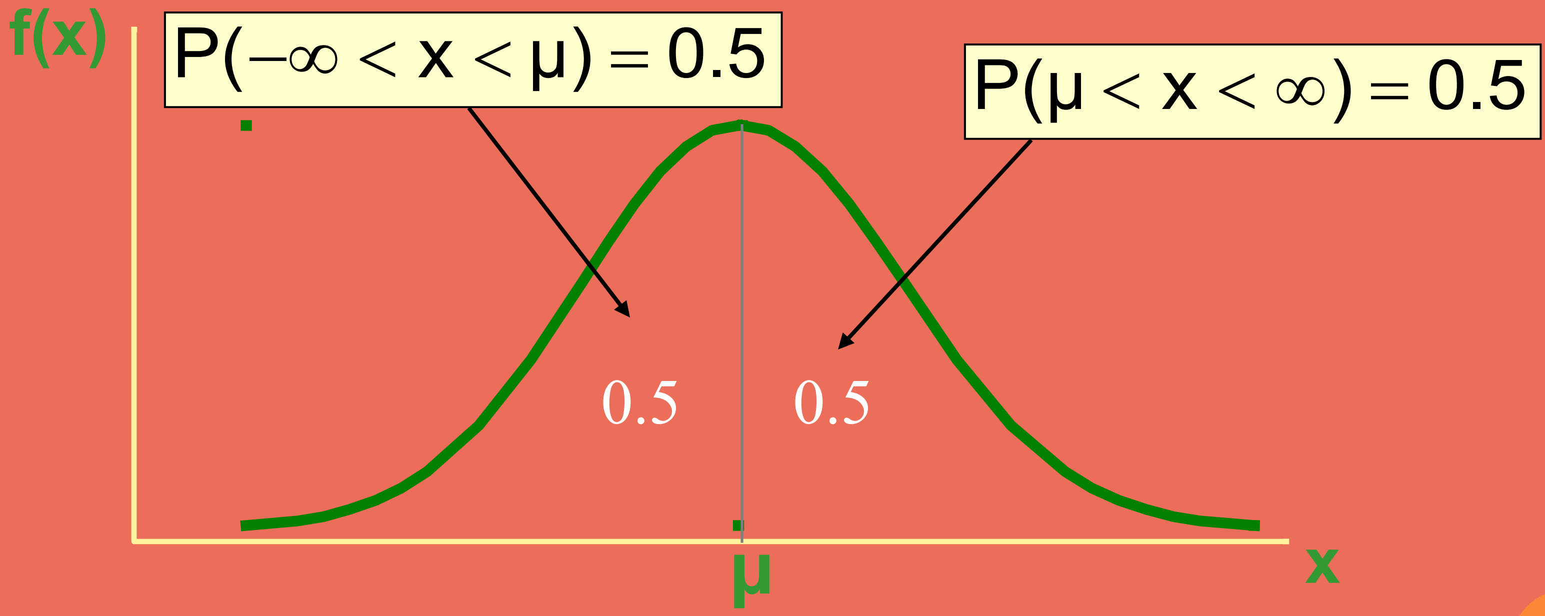
FINDING NORMAL PROBABILITIES

Probability is measured by the area under the curve



PROBABILITY AS AREA UNDER THE CURVE

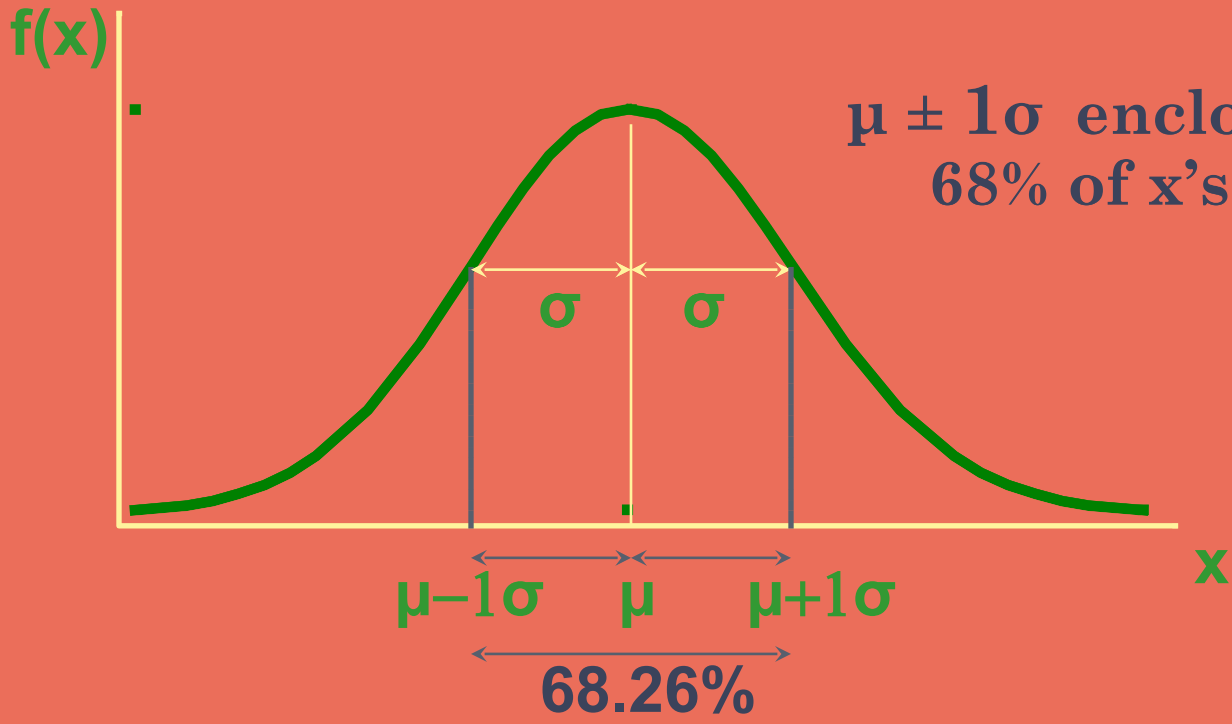
The total area under the curve is 1.0, and the curve is symmetric, so half is above the mean, half is below



$$P(-\infty < x < \infty) = 1.0$$

EMPIRICAL RULES

What can we say about the distribution of values around the mean? There are some general rules:

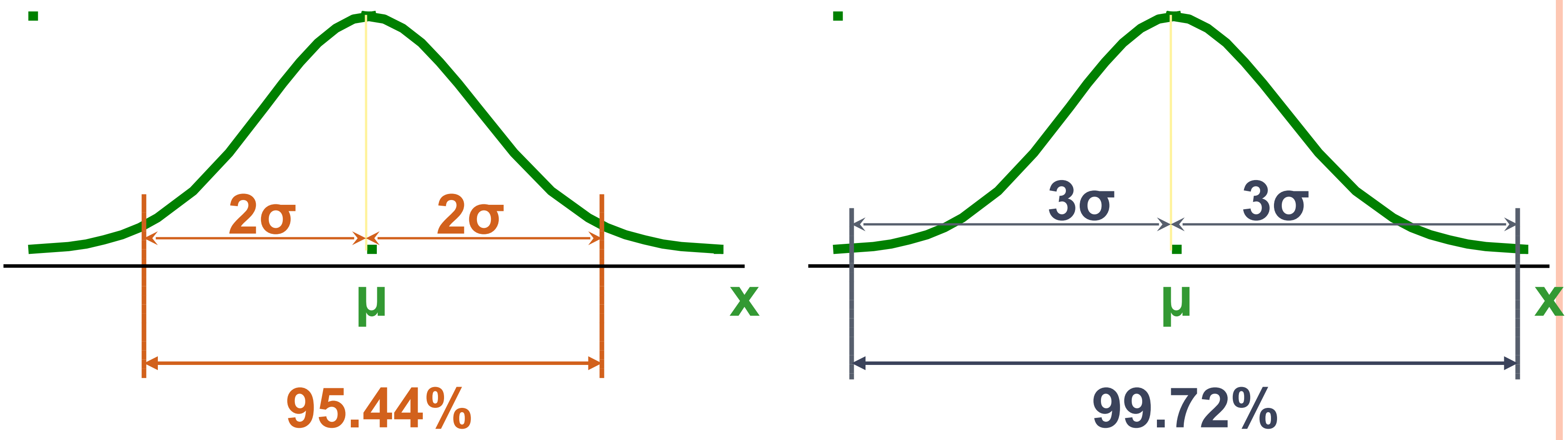


μ ± 1σ encloses about 68% of x's

THE EMPIRICAL RULE

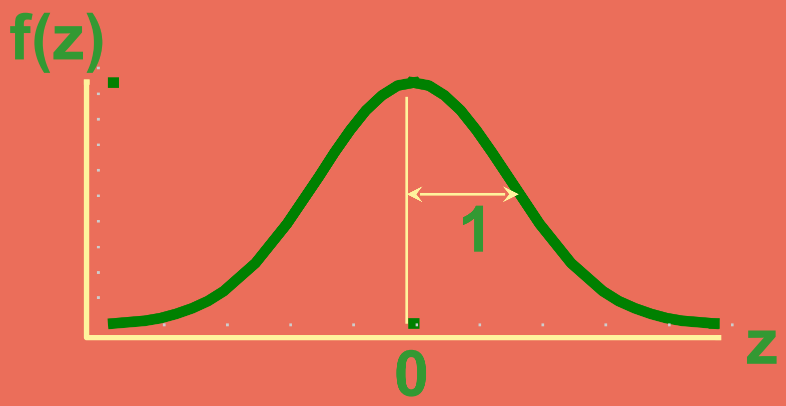
(continued)

- $\mu \pm 2\sigma$ covers about 95% of x's
- $\mu \pm 3\sigma$ covers about 99.7% of x's



THE STANDARD NORMAL DISTRIBUTION

- Also known as the “z” distribution
- Mean is defined to be 0
- Standard Deviation is 1



Values above the mean have positive z-values, values below the mean have negative z-values

THE STANDARD NORMAL

- Any normal distribution (with any mean and standard deviation combination) can be transformed into the standard normal distribution (z)
- Need to transform x units into z units

TRANSLATION TO THE STANDARD NORMAL DISTRIBUTION

- Translate from x to the standard normal (the “z” distribution) by subtracting the mean of x and dividing by its standard deviation:

$$z = \frac{x - \mu}{\sigma}$$

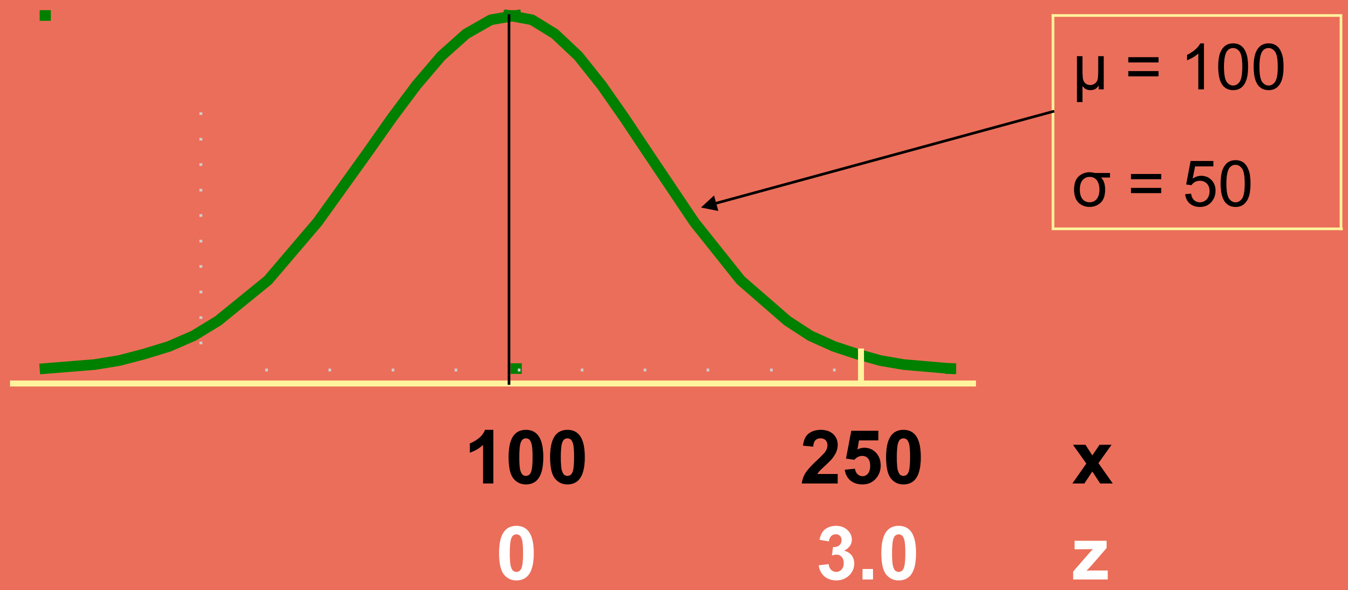
EXAMPLE

- If x is distributed normally with mean of 100 and standard deviation of 50, the z value for x = 250 is

$$z = \frac{x - \mu}{\sigma} = \frac{250 - 100}{50} = 3.0$$

- This says that x = 250 is three standard deviations (3 increments of 50 units) above the mean of 100.

COMPARING X AND Z UNITS



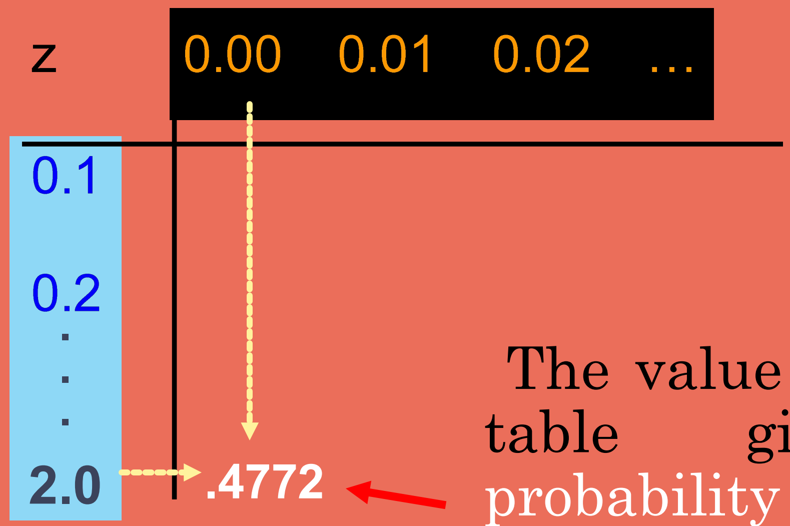
Note that the distribution is the same, only the scale has changed. We can express the problem in original units (x) or in standardized units (z)

THE STANDARD NORMAL TABLE

(continued)

The column gives the value of z to the second decimal point

The row shows the value of z to the first decimal point



The value within the table gives the probability from z = 0 up to the desired z value

P(0 < z < 2.00) = .4772

GENERAL PROCEDURE FOR FINDING PROBABILITIES

To find $P(a < x < b)$ when x is distributed normally:

- Draw the normal curve for the problem in terms of x
- Translate x-values to z-values
- Use the Standard Normal Table

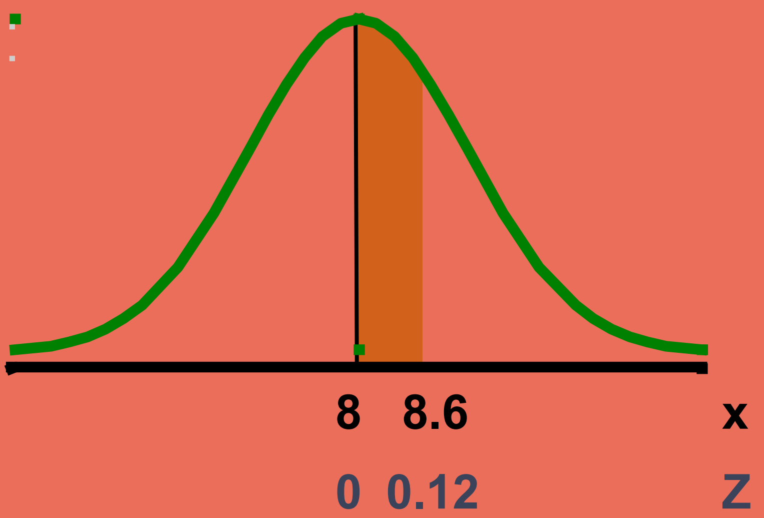
Z TABLE EXAMPLE

- Suppose x is normal with mean 8.0 and standard deviation 5.0. Find P(8 < x < 8.6)

Calculate z-values:

$$z = \frac{x - \mu}{\sigma} = \frac{8 - 8}{5} = 0$$

$$z = \frac{x - \mu}{\sigma} = \frac{8.6 - 8}{5} = 0.12$$

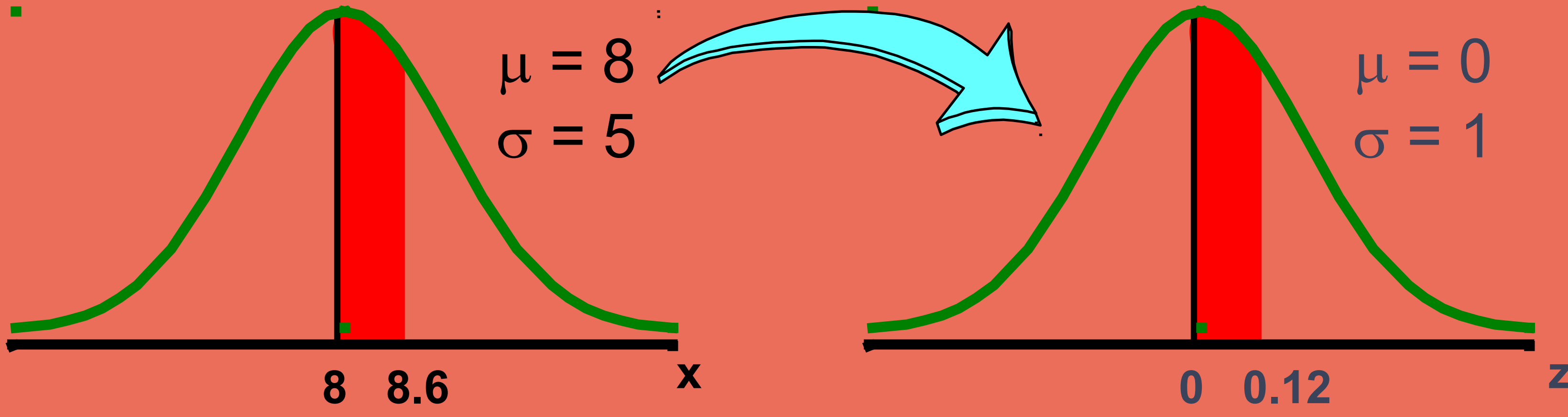


$$P(8 < x < 8.6)$$
$$= P(0 < z < 0.12)$$

Z TABLE EXAMPLE

(continued)

- Suppose x is normal with mean 8.0 and standard deviation 5.0. Find P(8 < x < 8.6)



P(8 < x < 8.6)

P(0 < z < 0.12)

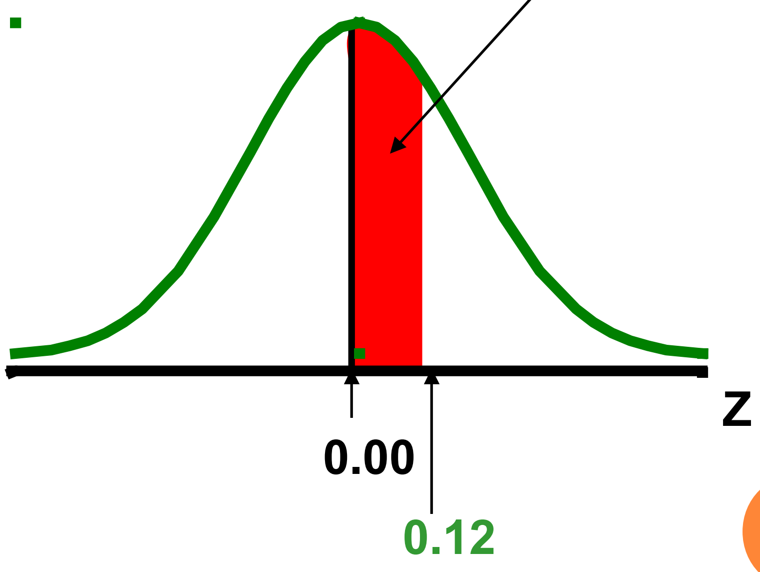
SOLUTION: FINDING P(0 < Z < 0.12)

Standard Normal Probability Table (Portion)

z	.00	.01	.02
0.0	.0000	.0040	.0080
0.1	.0398	.0438	.0478
0.2	.0793	.0832	.0871
0.3	.1179	.1217	.1255

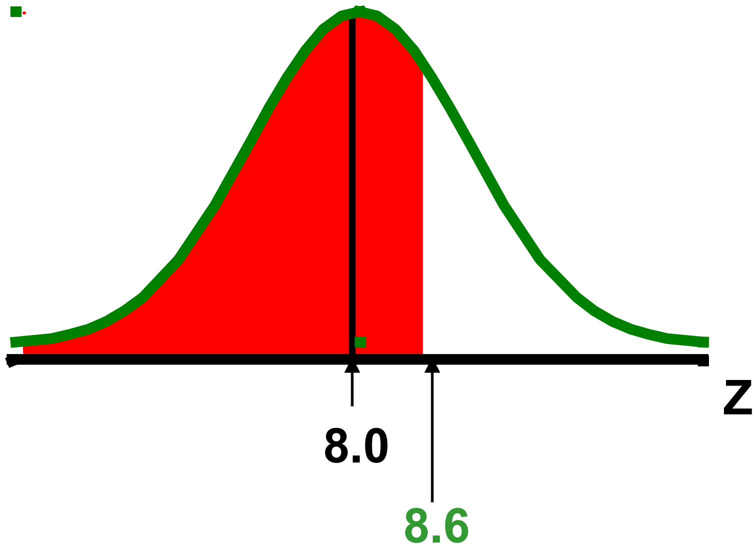
$P(8 < x < 8.6)$
$= P(0 < z < 0.12)$

.0478



FINDING NORMAL PROBABILITIES

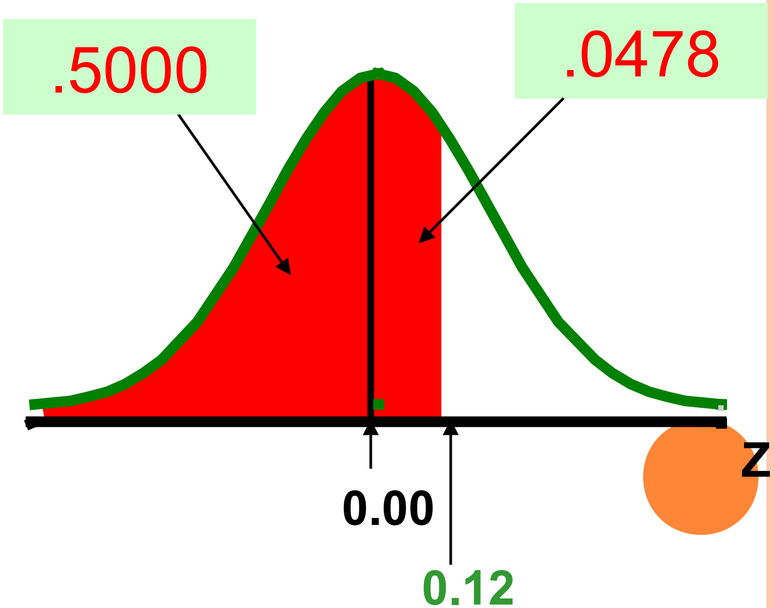
- Suppose x is normal with mean 8.0 and standard deviation 5.0.
- Now Find P(x < 8.6)



FINDING NORMAL PROBABILITIES *(continued)*

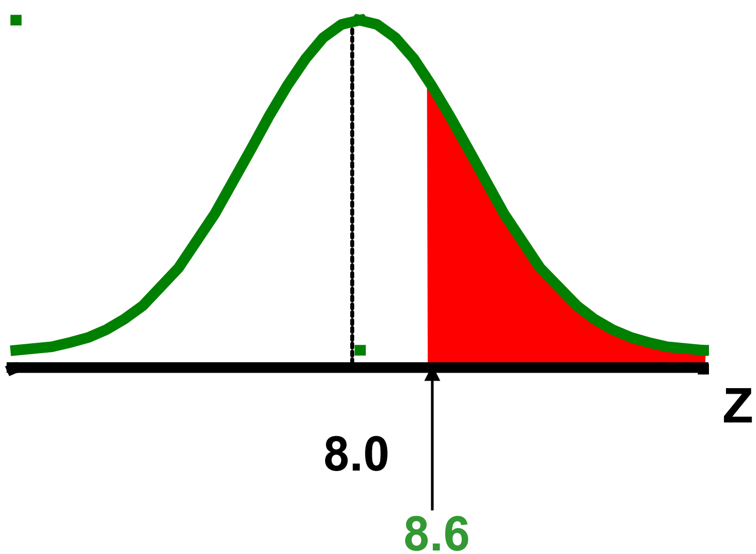
- Suppose x is normal with mean 8.0 and standard deviation 5.0.
- Now Find $P(x < 8.6)$

$$\begin{aligned} &P(x < 8.6) \\ &= P(z < 0.12) \\ &= P(z < 0) + P(0 < z < 0.12) \\ &= .5 + .0478 = .5478 \end{aligned}$$



UPPER TAIL PROBABILITIES

- Suppose x is normal with mean 8.0 and standard deviation 5.0.
- Now Find $P(x > 8.6)$



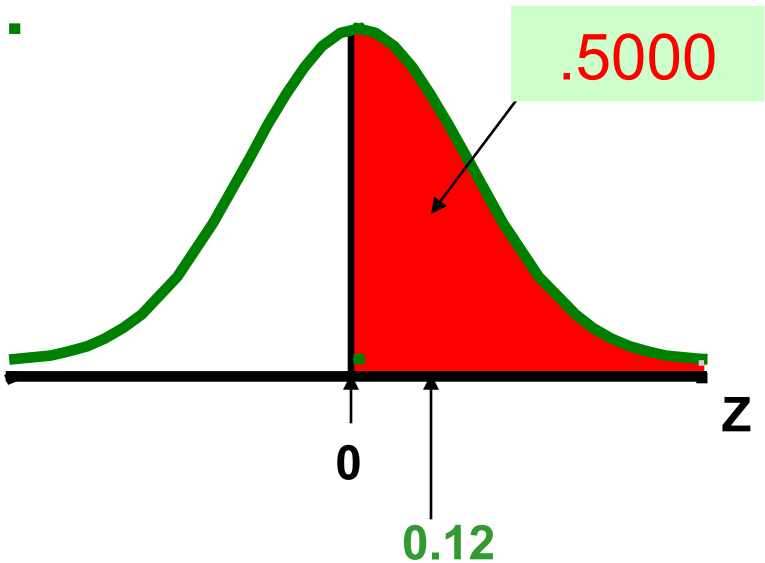
UPPER TAIL PROBABILITIES

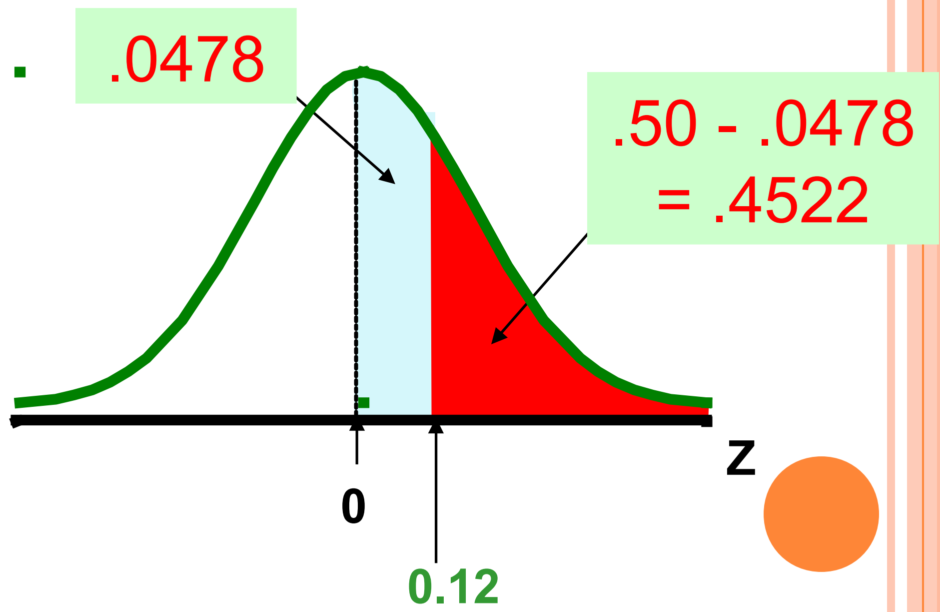
(continued)

- Now Find P(x > 8.6)…

$$P(x > 8.6) = P(z > 0.12) = P(z > 0) - P(0 < z < 0.12)$$

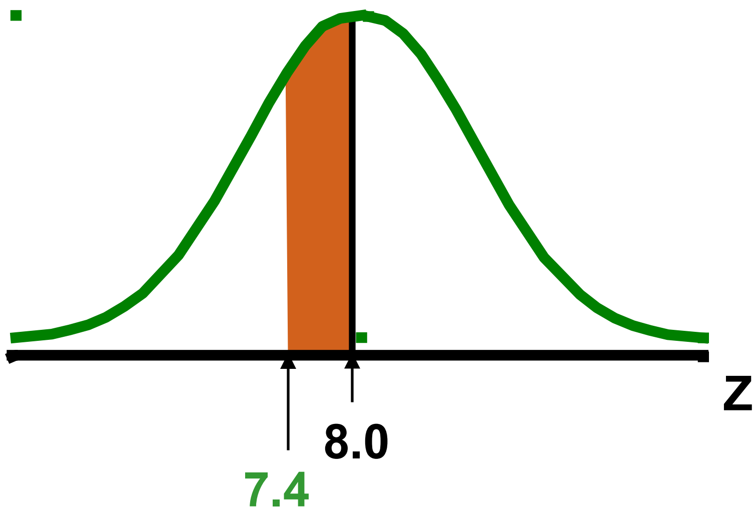
$$= .5 - .0478 = .4522$$





LOWER TAIL PROBABILITIES

- Suppose x is normal with mean 8.0 and standard deviation 5.0.
- Now Find $P(7.4 < x < 8)$



LOWER TAIL PROBABILITIES

(continued)

Now Find P(7.4 < x < 8)…

The Normal distribution is symmetric, so we use the same table even if z-values are negative:

$P(7.4 < x < 8)$

$= P(-0.12 < z < 0)$

$= .0478$