A Course In Business Statistics

4th Edition

Chapter 5

Discrete and Continuous Probability Distributions

Chapter Goals

After completing this chapter, you should be able to:

- Apply the binomial distribution to applied problems
- Compute probabilities for the Poisson and hypergeometric distributions
- Find probabilities using a normal distribution table and apply the normal distribution to business problems
- Recognize when to apply the uniform and exponential distributions

Probability Distributions

- **Probability Distributions**
  - **Discrete** Probability Distributions
    - Binomial
    - Poisson
    - Hypergeometric
  - **Continuous** Probability Distributions
    - Normal
    - Uniform
    - Exponential

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

Probability Distributions

- **Discrete** Probability Distributions
  - Binomial
  - Poisson
  - Hypergeometric

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

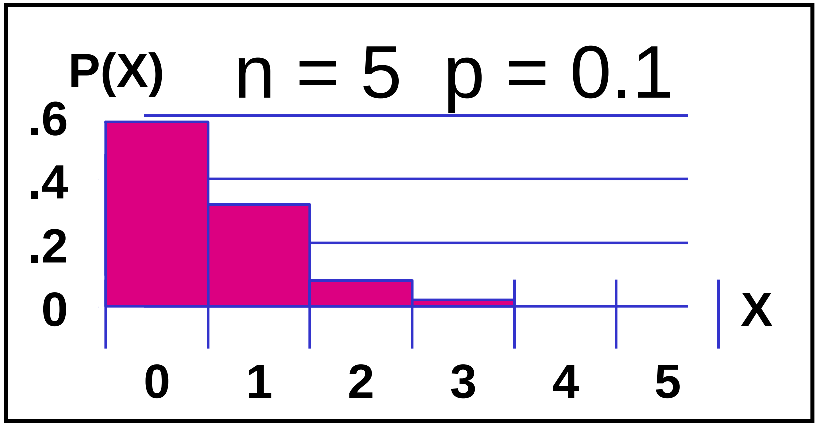
q = (1 - .5) = .5

x = 0, 1, 2, 3, 4

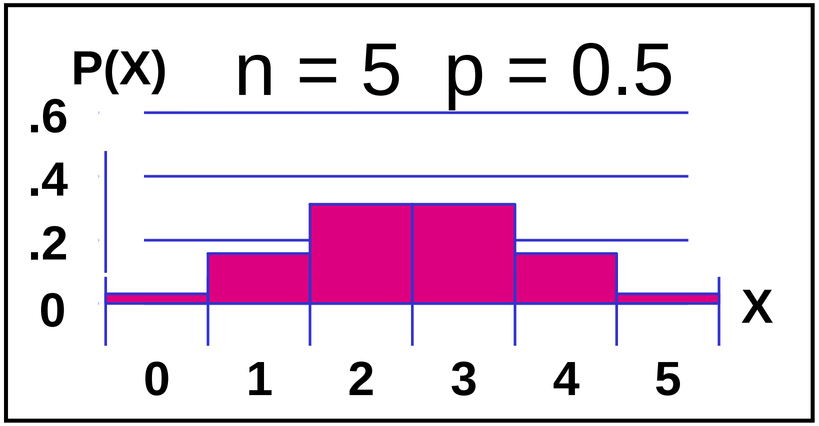
Binomial Distribution

- 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

$$\sigma^2 = npq$$

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

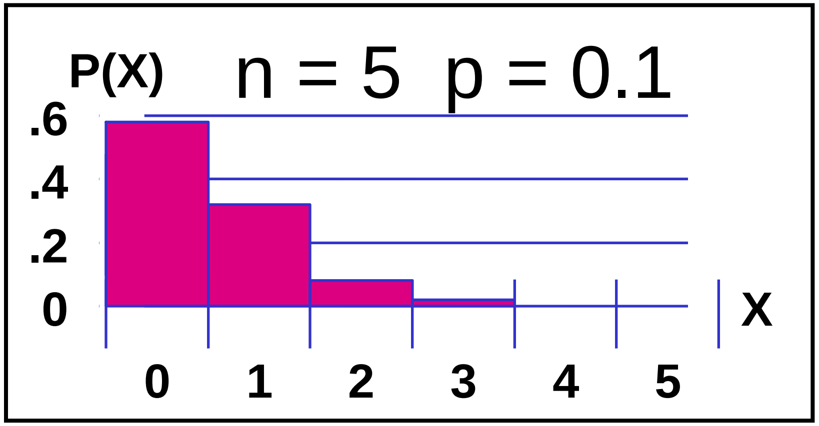
Where n = sample size
p = probability of success
q = (1 – p) = probability of failure

Binomial Characteristics

Examples

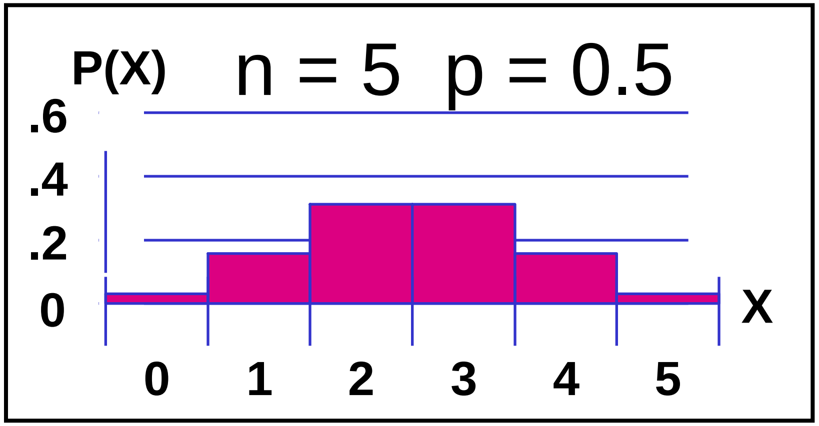
$$\mu = np = (5)(.1) = 0.5$$

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



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

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



The Hypergeometric Distribution

Probability Distributions

Discrete Probability Distributions

- Binomial
- Poisson
- Hypergeometric

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)

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

Where

N = Population size

S = 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
S = 4	x = 2

$$P(x=2) = \frac{C_2^4 . C_{3-2}^{10-4}}{C_3^{10}} = \frac{C_1^6 \; C_2^4}{C_3^{10}} = \frac{(6)(6)}{120} = 0.3$$

The Poisson Distribution

- **Probability Distributions**
  - **Discrete** Probability Distributions
    - Binomial
    - Poisson
    - Hypergeometric

The Poisson Distribution

- Uses of Poisson Distribution:
  - First proposed by Simeon Poisson (1781-1840)
  - The number of failures in a large computer system during a given day
  - The number of replacement orders for a part received by a firm in a given month
  - The number of ships arriving at a loading facility during a 6-hour loading period
  - The number of delivery trucks to arrive at a central warehouse in an hour

The Poisson Distribution

- Uses of Poisson Distribution:
  - The number of dents, scrathes, or other defects in a large roll of sheet metal used to manufacture filters
  - The number of customers to arrive for flights during each 15 min time interval from 3pm to 6pm on weekdays
  - The number of customers to arrive at a checkout aisle in your local grocery store during a particular time interval

The Poisson Distribution

- We can use the Poisson distribution to determine the probability of each of these random variables, which are characterized as the number of occurences or successes of a certain event in a given continuous interval (such as time or space – area, length).

The Poisson Distribution

- Assumptions 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 and it is **constant**.
  - 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

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

where:

x = the number of successes over a given time or space

λ = the expected number of successes per time or space unit, $\lambda > 0$

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

Poisson Distribution Characteristics

- Mean

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

- Variance and Standard Deviation

$$\sigma^2 = \lambda$$

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

Example 1 for Poisson Distrbution

- Andrew Whittaker, computer centre manager, reports that his computer system experienced 3 component failures during the past 100 days.
  - a) What is the probability of no failures in a given day?
  - b) What is the probability of one or more component failures in a given day?
  - c) What is the probability of at least 2 failures in a 3-day period?

Example 2 for Poisson Distrbution

- Customers arrive at a photocopying machine at an average rate of two every 5 mins. Assume that these arrivals are independent, with a constant arrival rate, and that this problem follows a Poisson model, with X denoting the number of arriving customers at a 5min period and mean $\lambda=2$. Find the probability that more than two customers arrive in a 5min period?

The Normal Distribution

Probability Distributions

- **Continuous** Probability Distributions
  - Normal
  - Uniform
  - Exponential

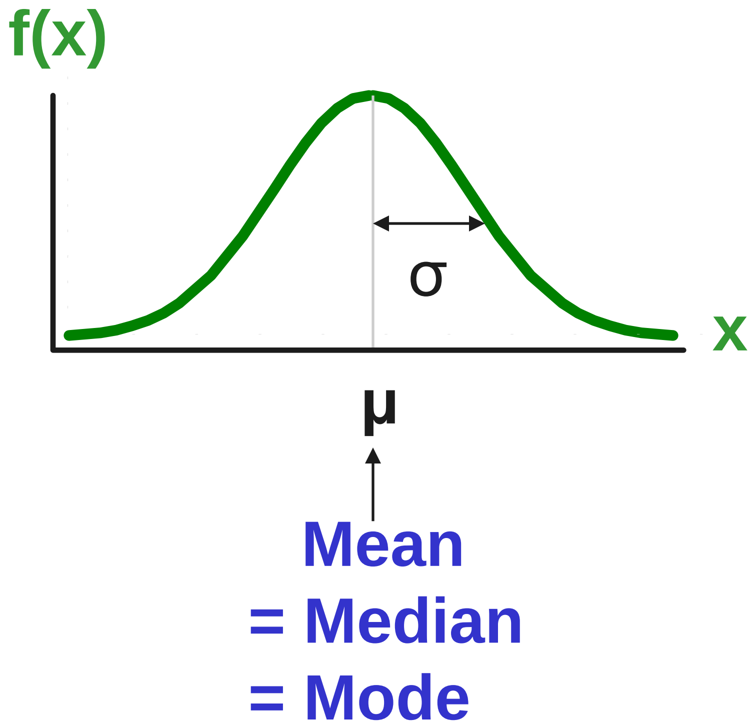
The Normal Distribution

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

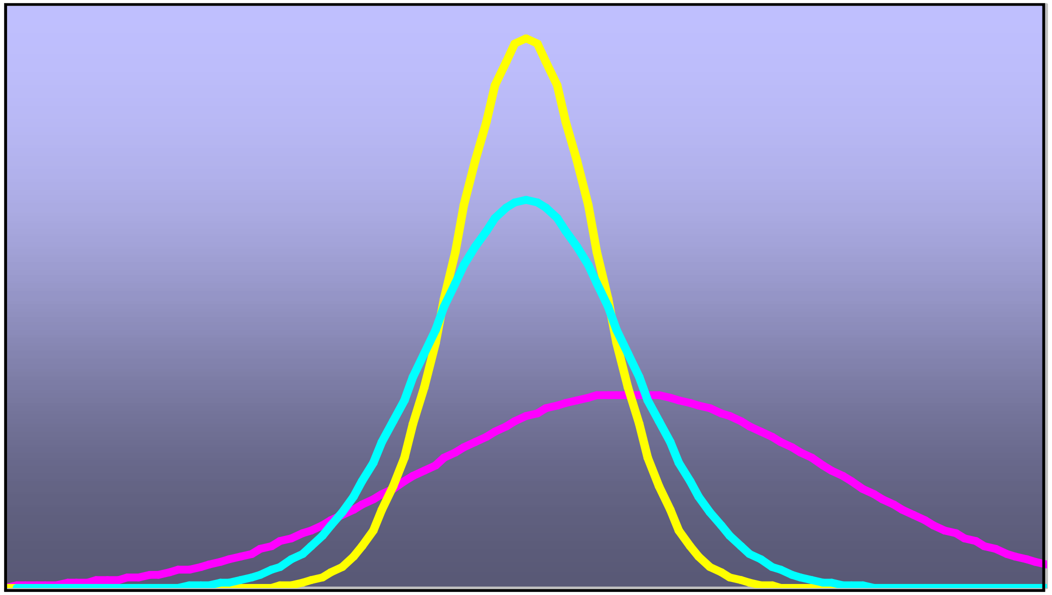
Location is determined by the mean, μ

Spread is determined by the standard deviation, σ

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

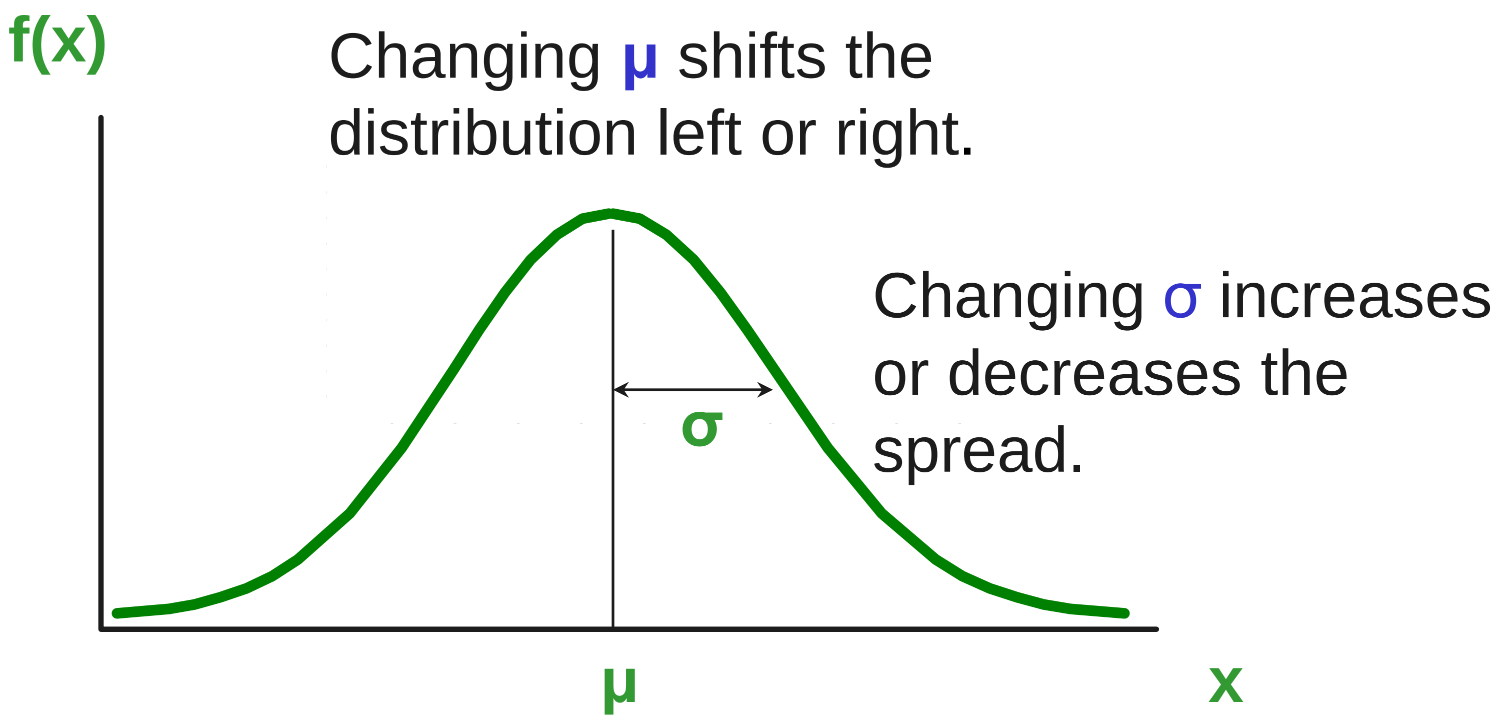


Many Normal Distributions



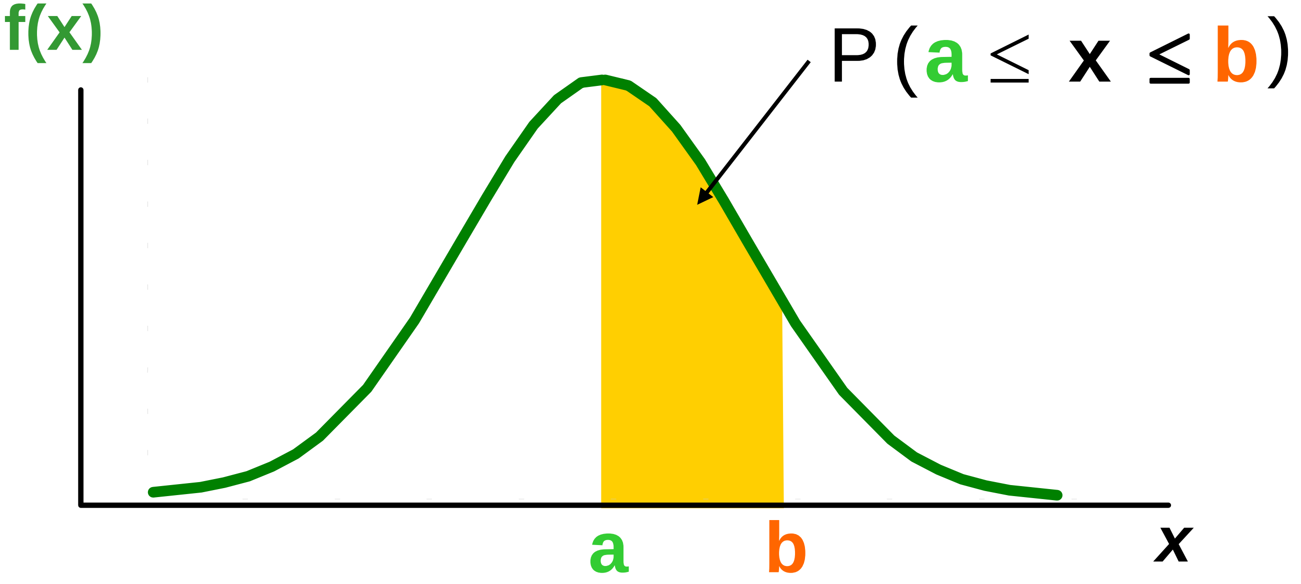
By varying the parameters μ and σ, we obtain different normal distributions

The Normal Distribution Shape



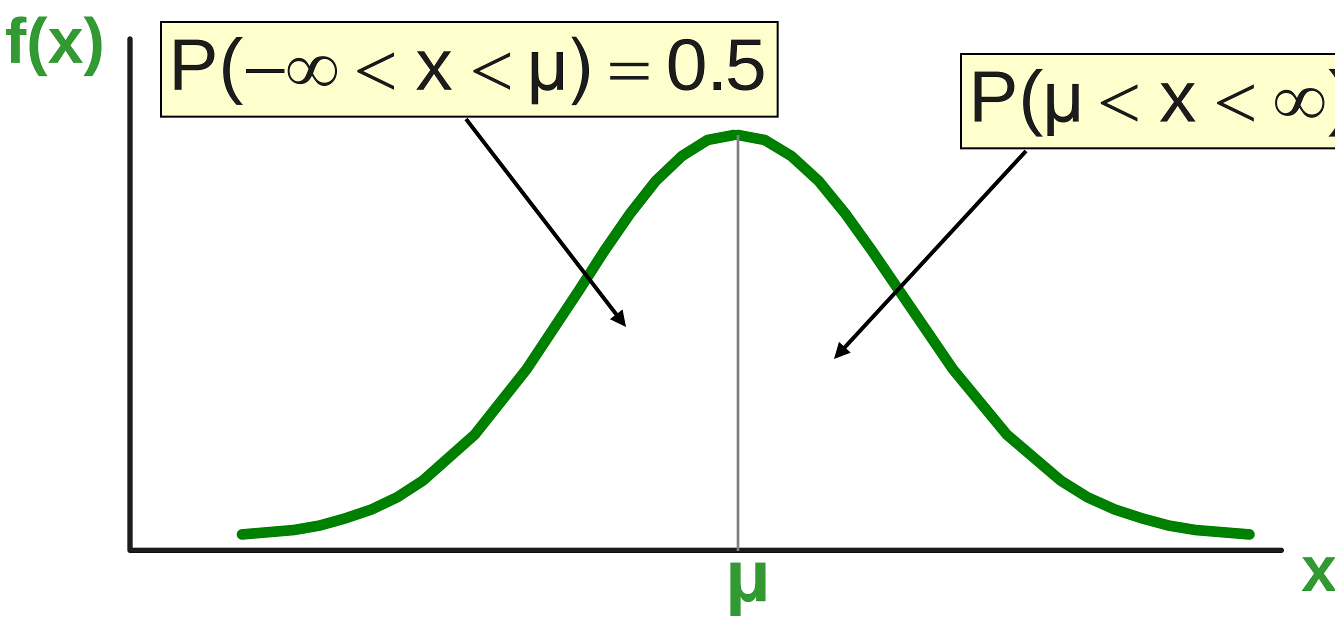
Finding Normal Probabilities

Probability is measured by the area under the curve



Probability as Area Under the Curve

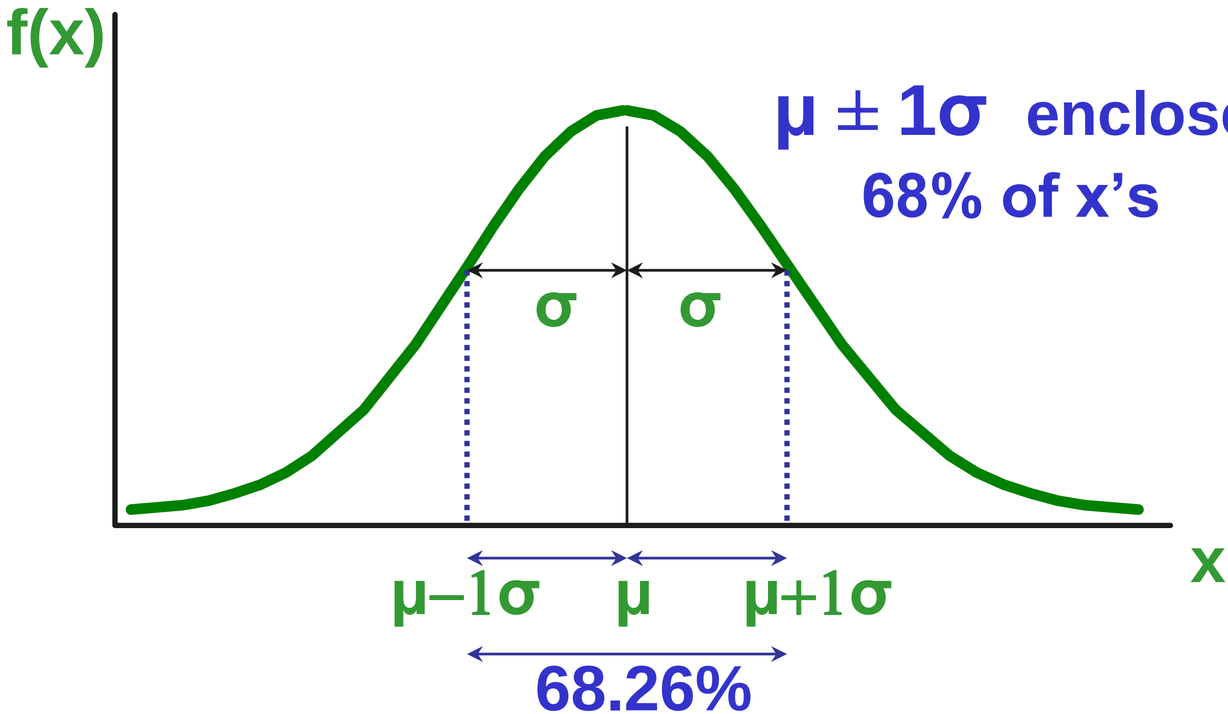
The total area under the curve is 1.0, and the curve is symmetric, so half is above the mean, half is below



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

Empirical Rules

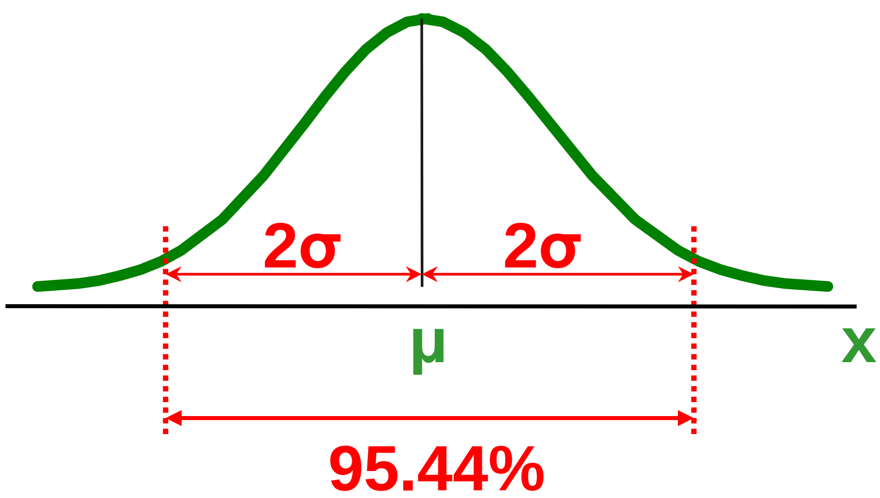
What can we say about the distribution of values around the mean? There are some general rules:

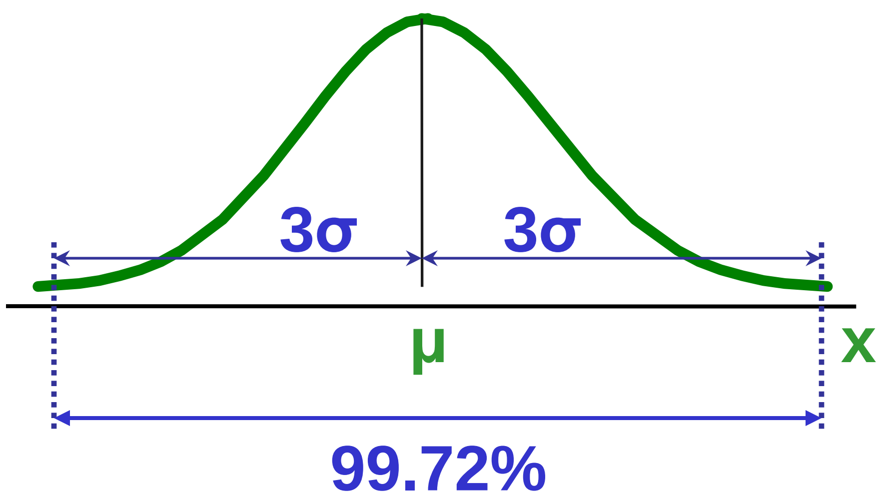


The Empirical Rule

(continued)

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





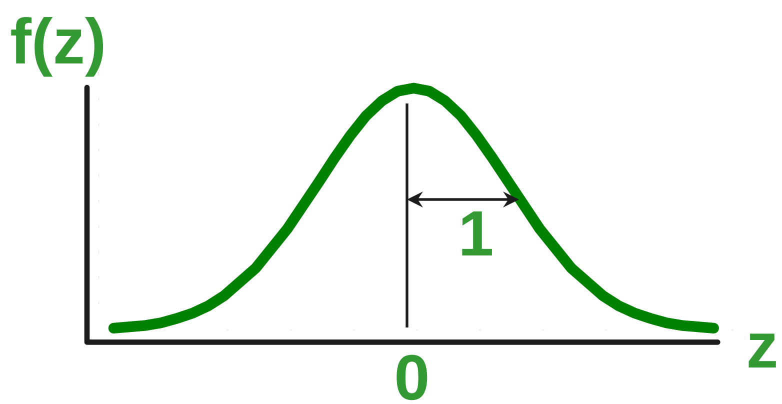
Importance of the Rule

- If a value is about 2 or more standard deviations away from the mean in a normal distribution, then it is far from the mean

- The chance that a value that far or farther away from the mean is highly unlikely, given that particular mean and standard deviation

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:

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

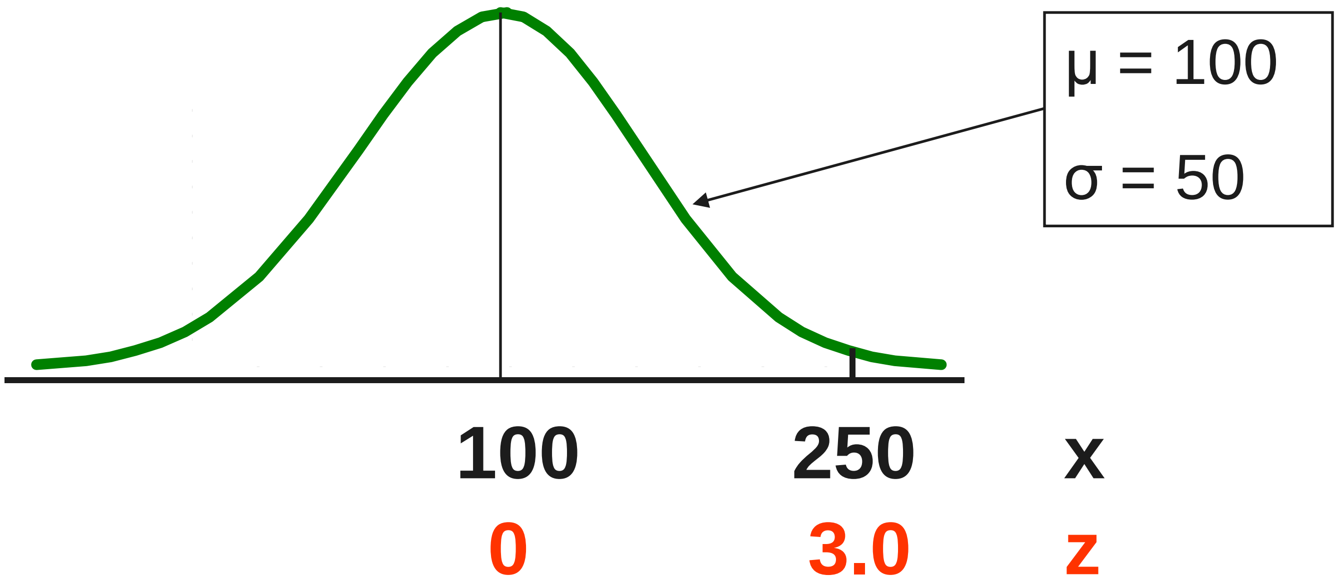
Example

- 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)

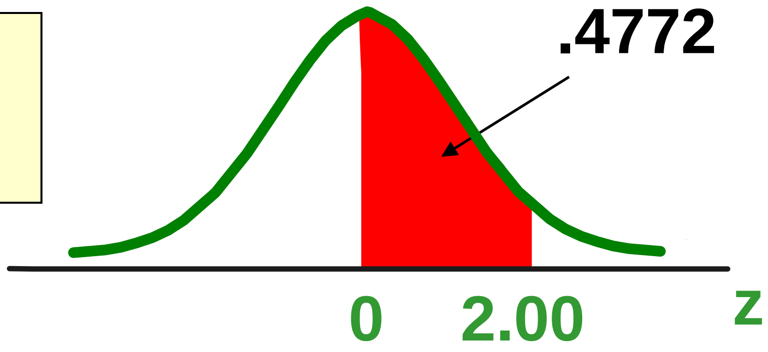
The Standard Normal Table

- The Standard Normal table in the textbook (Appendix D)

  gives the probability from the mean (zero) up to a desired value for z

Example:

P(0 < z < 2.00) = .4772



The Standard Normal Table

(continued)

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

z	0.00	0.01	0.02	…
0.1				
0.2				
.				
.				
.				
2.0	.4772			

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

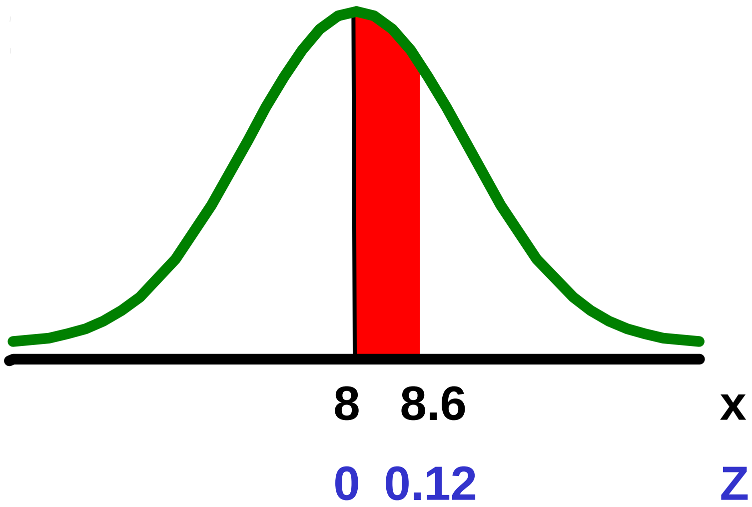
Z Table example

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

Calculate z-values:

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

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



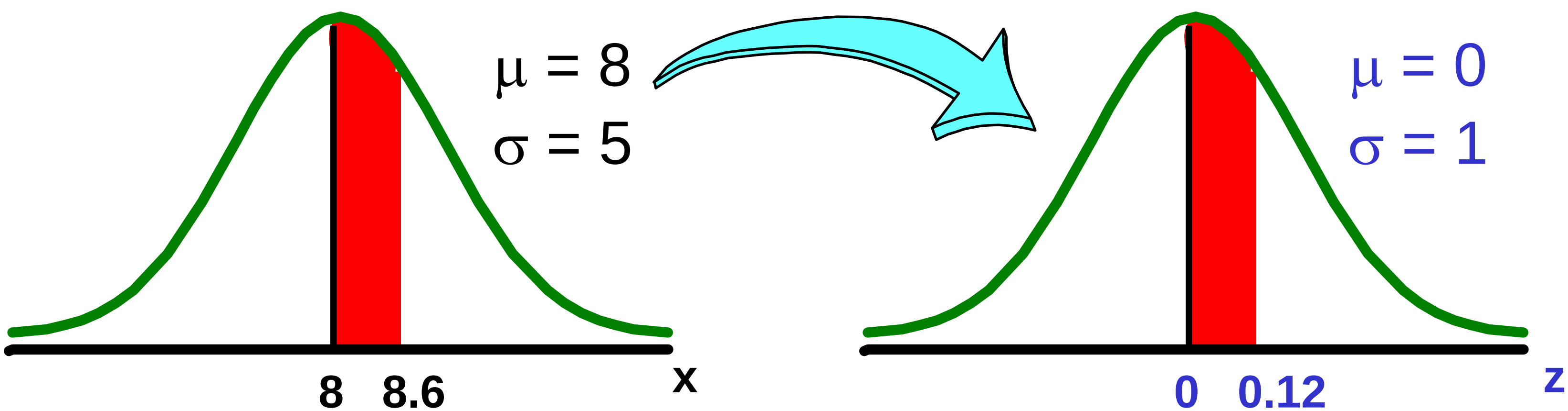
P(8 < x < 8.6)

= P(0 < z < 0.12)

Z Table example

(continued)

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



$P(8 < x < 8.6)$ $\quad$ $P(0 < z < 0.12)$

Solution: Finding P(0 < z < 0.12)

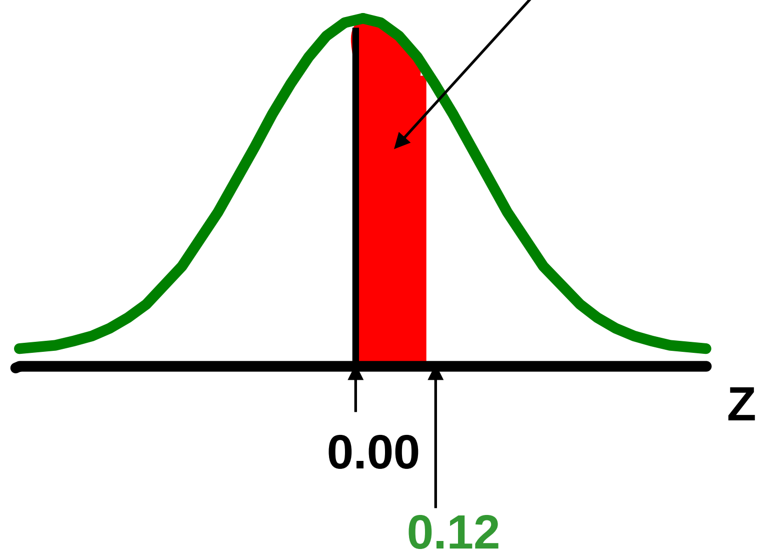
Standard Normal Probability Table (Portion)

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

$P(8 < x < 8.6)$

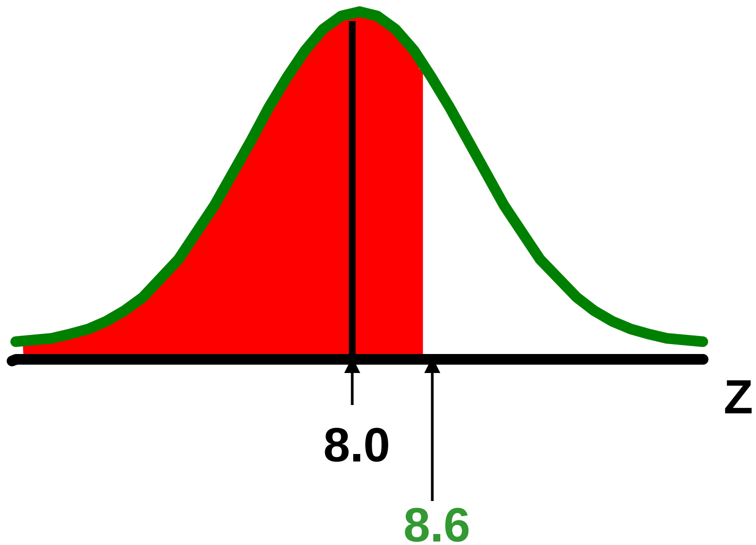
$= P(0 < z < 0.12)$

.0478



Finding Normal Probabilities

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

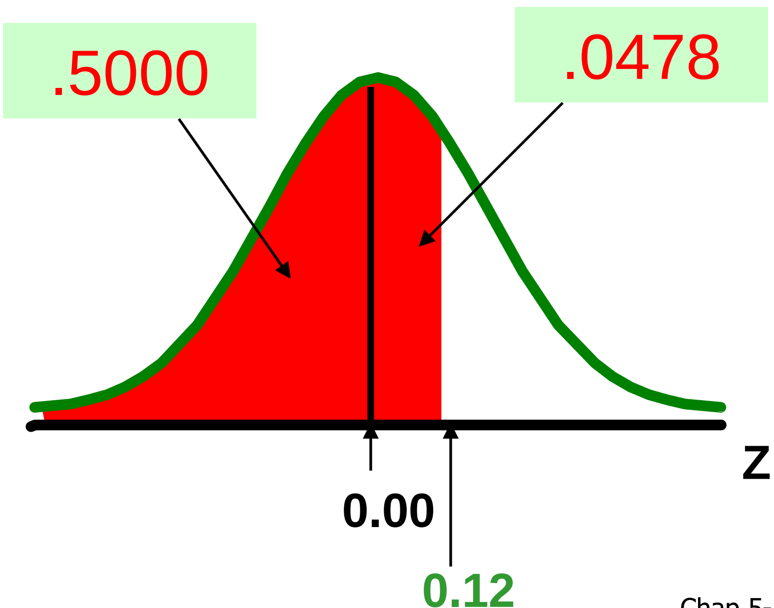


Finding Normal Probabilities

(continued)

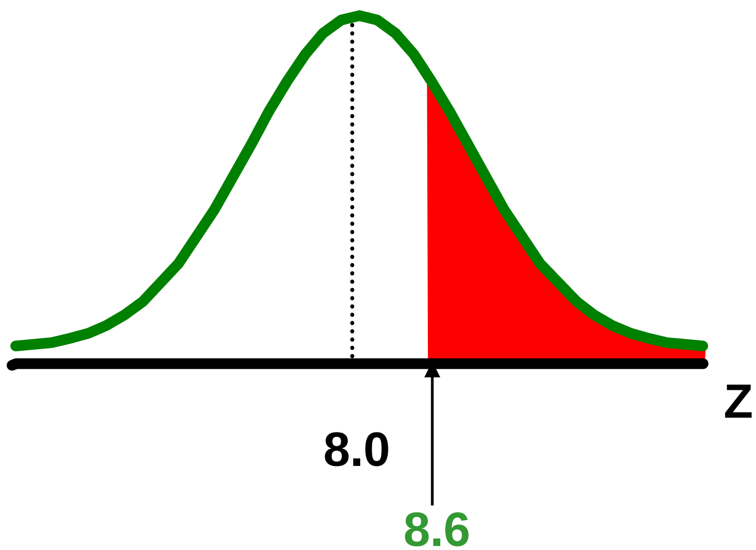
- Suppose x is normal with mean 8.0 and standard deviation 5.0.
- Now Find P(x < 8.6)

$$P(x < 8.6)$$
$$= P(z < 0.12)$$
$$= P(z < 0) + P(0 < z < 0.12)$$
$$= .5 + .0478 = \boxed{.5478}$$



Upper Tail Probabilities

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



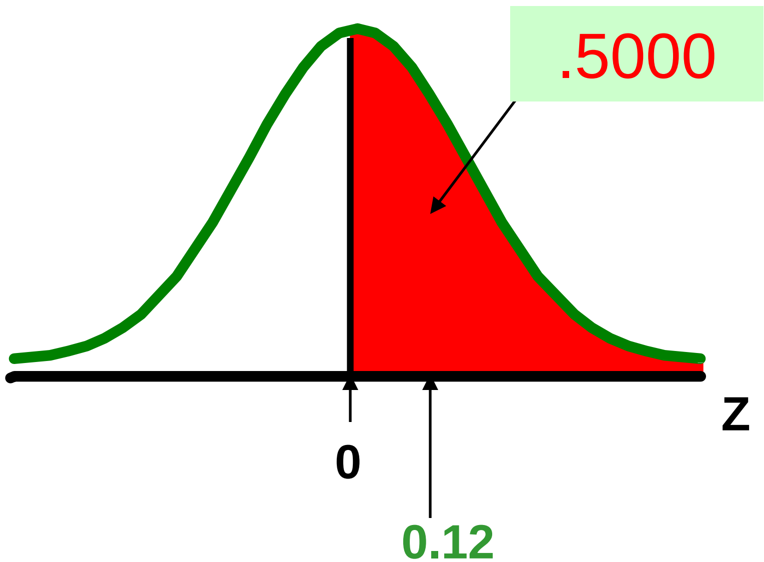
Upper Tail Probabilities

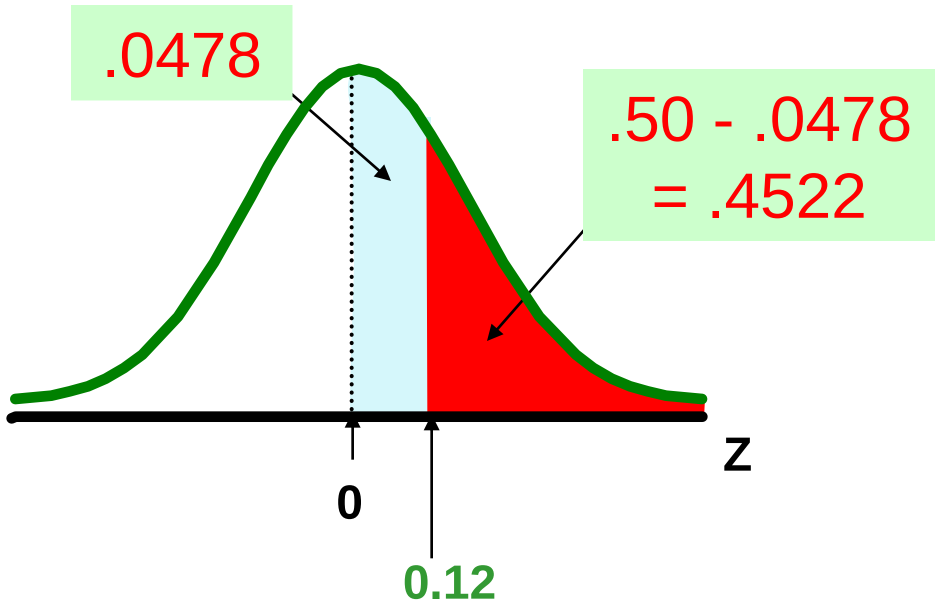
(continued)

- Now Find P(x > 8.6)…

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

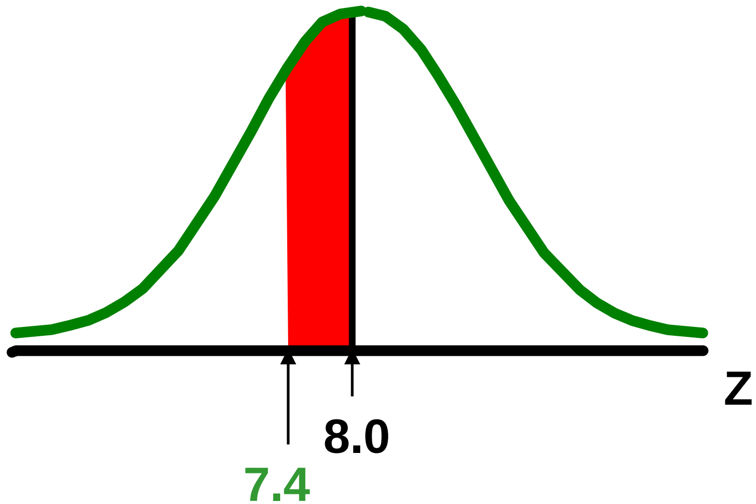
$$= .5 - .0478 = \boxed{.4522}$$





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)

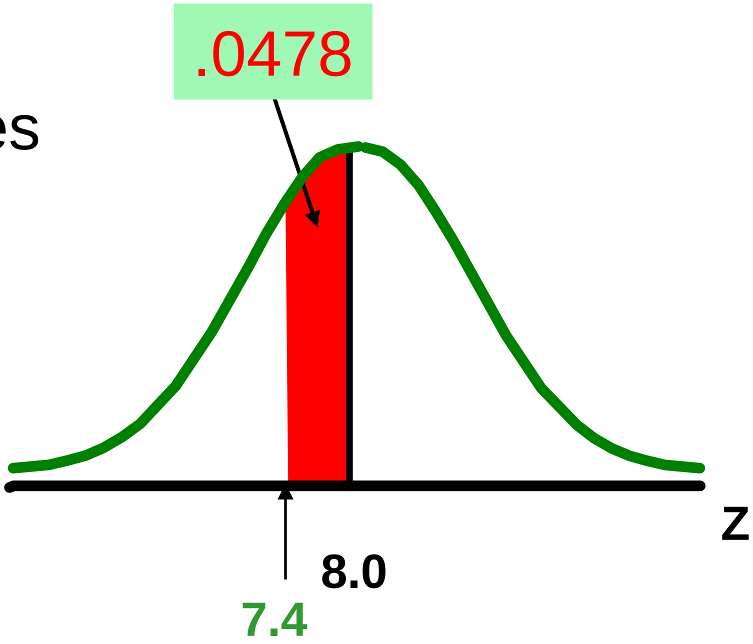
Now Find P(7.4 < x < 8)…

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

$P(7.4 < x < 8)$

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

$= \boxed{.0478}$



The Exponential Distribution

- **Probability Distributions**
  - **Continuous** Probability Distributions
    - Normal
    - Uniform
    - Exponential

The Exponential Distribution

- Used to measure the time that elapses between two occurrences of an event (the time between arrivals)

  - Examples:
    - Time between trucks arriving at an unloading dock
    - Time between transactions at an ATM Machine
    - Time between phone calls to the main operator

The Exponential Distribution

- The probability that an arrival time is equal to or less than some specified time a is

$$P(0 \le x \le a) = 1 - e^{-\lambda a}$$

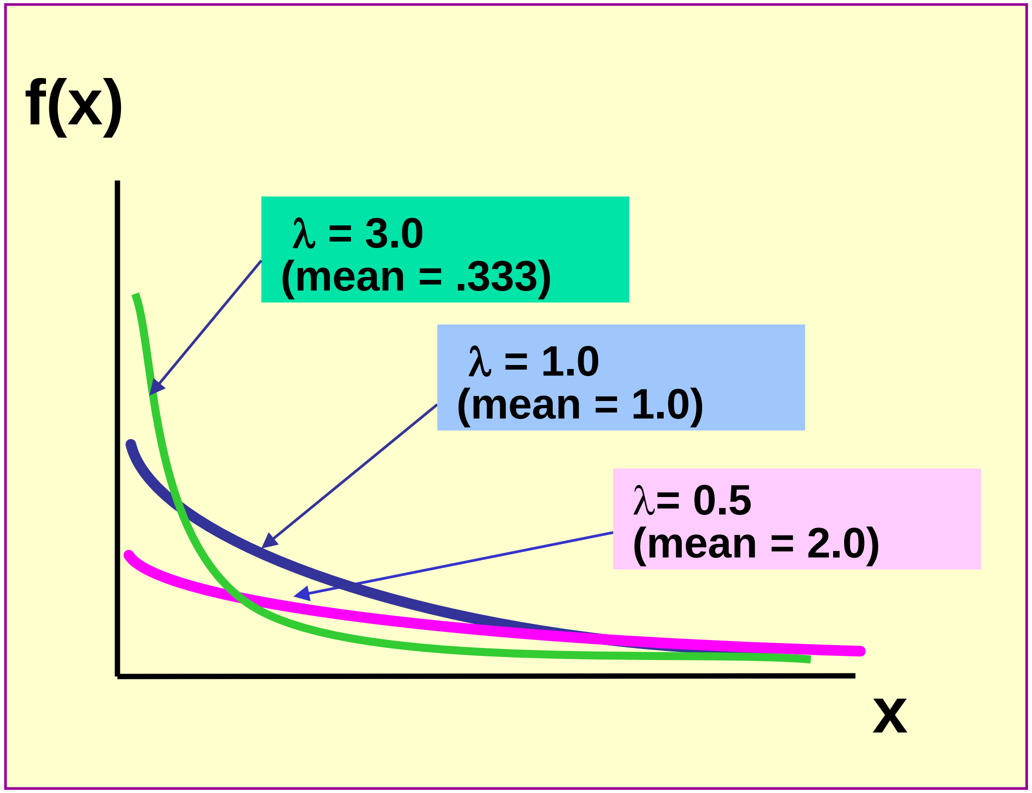
where $1/\lambda$ is the mean time between events

Note that if the number of occurrences per time period is Poisson with mean λ, then the time between occurrences is exponential with mean time $1/\lambda$

Exponential Distribution

(continued)

- Shape of the exponential distribution



Example

Example: Customers arrive at the claims counter at the rate of 15 per hour (Poisson distributed). What is the probability that the arrival time between consecutive customers is less than five minutes?

- Time between arrivals is exponentially distributed with mean time between arrivals of 4 minutes (15 per 60 minutes, on average)
- $1/\lambda = 4.0$, so $\lambda = .25$
- $P(x < 5) = 1 - e^{-\lambda a} = 1 - e^{-(.25)(5)} = \boxed{.7135}$

Chapter Summary

- Reviewed key discrete distributions
  - binomial, poisson, hypergeometric
- Reviewed key continuous distributions
  - normal, uniform, exponential
- Found probabilities using formulas and tables
- Recognized when to apply different distributions
- Applied distributions to decision problems