Chapter 6 Continuous Random Variables

In Chapter 5 we defined discrete and continuous random variables. We also discussed discrete probability distributions, which are used to compute the probabilities of values of discrete random variables. In this chapter we discuss **continuous probability distributions.** These are used to find probabilities concerning continuous random variables. We begin by explaining the general idea behind a continuous probability distribution. Then we present three important continuous distributions—**the uniform, normal,** and **exponential distributions.** We also study when and how the normal distribution can be used to approximate the binomial distribution (which was discussed in Chapter 5).

We will illustrate the concepts in this chapter by using two cases:

The Car Mileage Case: A competitor claims that its midsize car gets better mileage than an automaker's new midsize model. The automaker uses sample information and a probability based on the normal distribution to provide strong evidence that the competitor's claim is false.

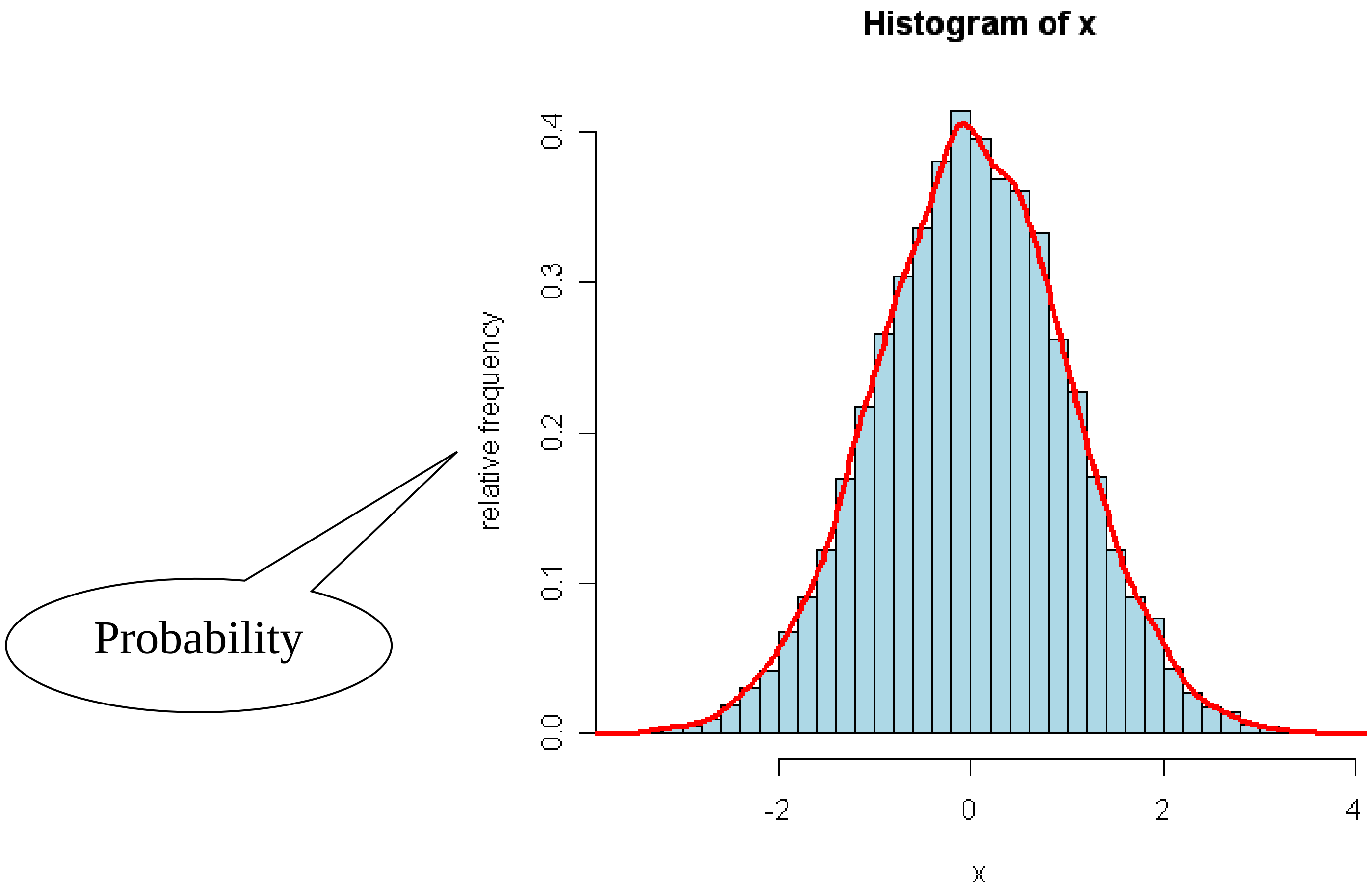
The Coffee Temperature Case: A fast-food restaurant uses the normal distribution to estimate the proportion of coffee it serves that has a temperature (in degrees Fahrenheit) inside the range 153° to 167°, the customer requirement for best-tasting coffee.

Outline

6.1 Continuous Probability Distributions
6.2 The Uniform Distribution
6.3 The Normal Probability Distribution
6.4 Approximating the Binomial Distribution by Using the Normal Distribution (Optional)
6.5 The Exponential Distribution (Optional)
6.6 The Normal Probability Plot (Optional)

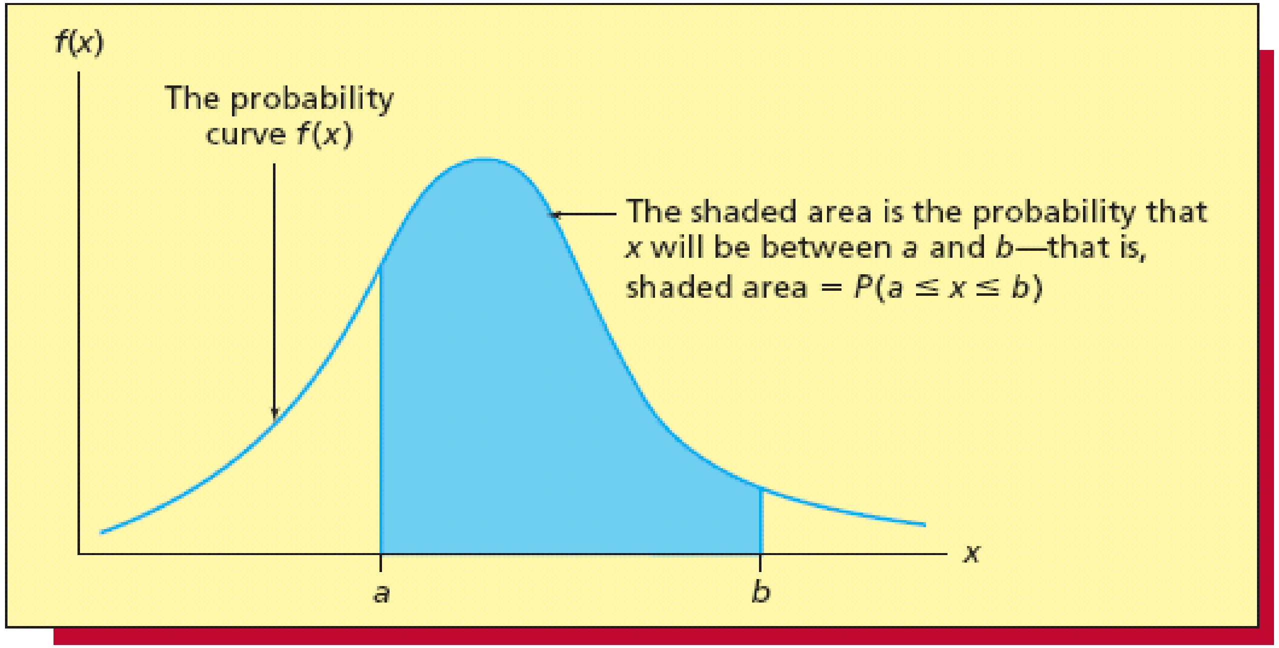
6.1 Continuous Probability Distributions

- A continuous random variable may assume any numerical value in one or more intervals
  - Car mileage
  - Temperature
- Use a continuous probability distribution to assign probabilities to intervals of values
- Uses a continuous probability distribution


Probability
Histogram of x
relative frequency
0.0
0.1
0.2
0.3
0.4
-2
0
2
4
x

Definition: The curve $f(x)$ is the continuous probability distribution of the continuous random variable x if the probability that x will be in a specified interval of numbers is the area under the curve $f(x)$ corresponding to the interval

- Other names for a continuous probability distribution:
  - probability curve, or
  - probability density function



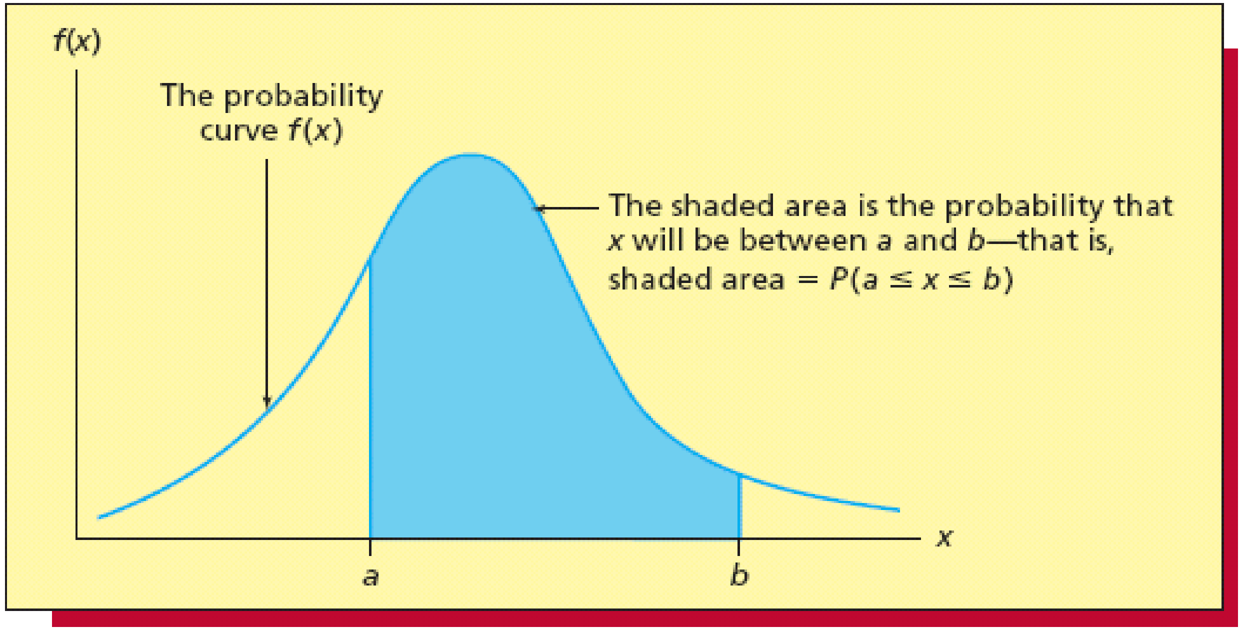
Properties of Continuous Probability Distributions

- Properties of $f(x)$: $f(x)$ is a continuous function such that
- 1. $f(x) \geq 0$ for all x
- 2. The total area under the curve of $f(x)$ is equal to 1

Essential point: *An area under a continuous probability distribution is a probability*

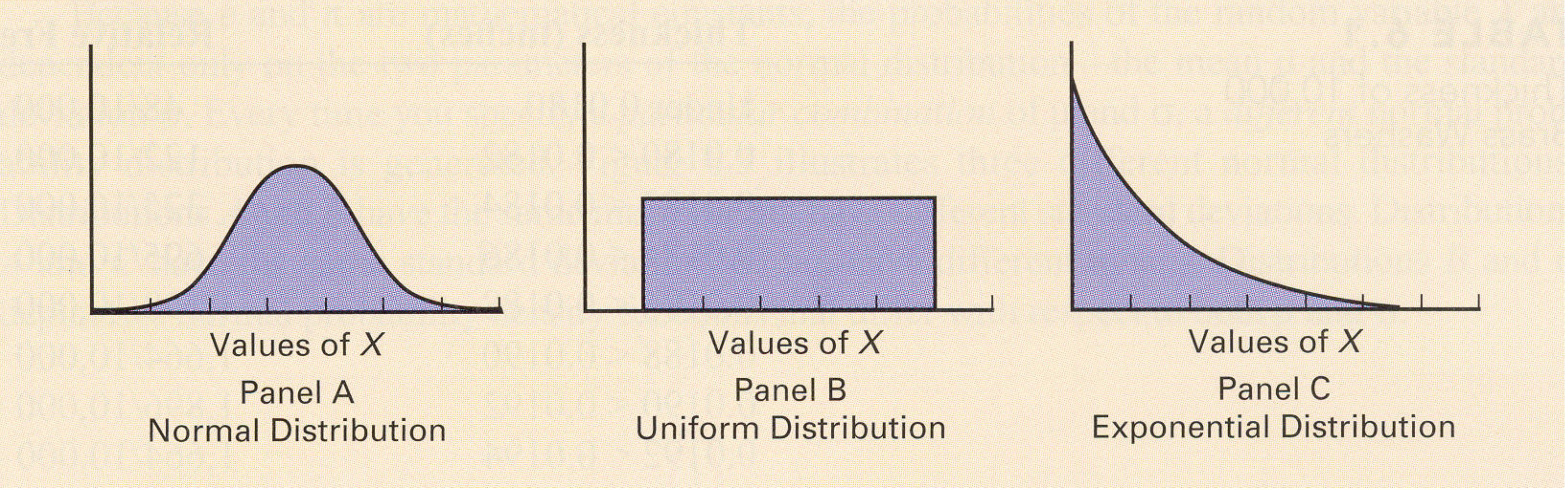
Area and Probability

- The blue-colored area under the probability curve *f(x)* from the value $x = a$ to $x = b$ is the probability that x could take any value in the range a to b
  - Symbolized as $P(a \leq x \leq b)$
    - Or as $P(a < x < b)$, because each of the interval endpoints has a probability of *0*



Continuous Probability Distributions

- For a continuous random variable, its probability density function gives a complete information about the variable.
- The most useful density functions are Normal distribution, uniform distribution and Exponetial distribution



We have seen that to calculate a probability concerning a continuous random variable, we must compute an appropriate area under the curve $f(x)$. Because there is no area under a continuous curve at a single point, or number, on the real line, the probability that a continuous random variable x will equal a single numerical value is always equal to 0. It follows that if $[a, b]$ denotes

an arbitrary interval of numbers on the real line, then $P(x = a) = 0$ and $P(x = b) = 0$. Therefore, $P(a \leq x \leq b)$ equals $P(a < x < b)$ because each of the interval endpoints a and b has a probability that is equal to 0.

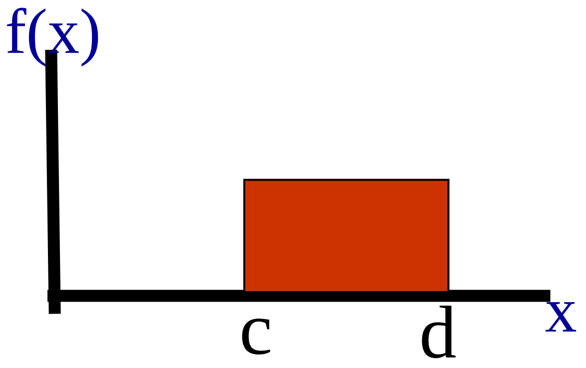
Distribution Shapes

- Symmetrical and rectangular
  - The uniform distribution
    - Section 6.2
- Symmetrical and bell-shaped
  - The normal distribution
    - Section 6.3
- Skewed
  - Skewed either left or right
    For Example, the right-skewed exponential distribution

6.2 The Uniform Distribution

If c and d are numbers on the real line ($c < d$), the probability curve describing the uniform distribution is

$$f(x) = \begin{cases} \dfrac{1}{d-c} & \text{for } c \le x \le d \\ 0 & \text{otherwise} \end{cases}$$

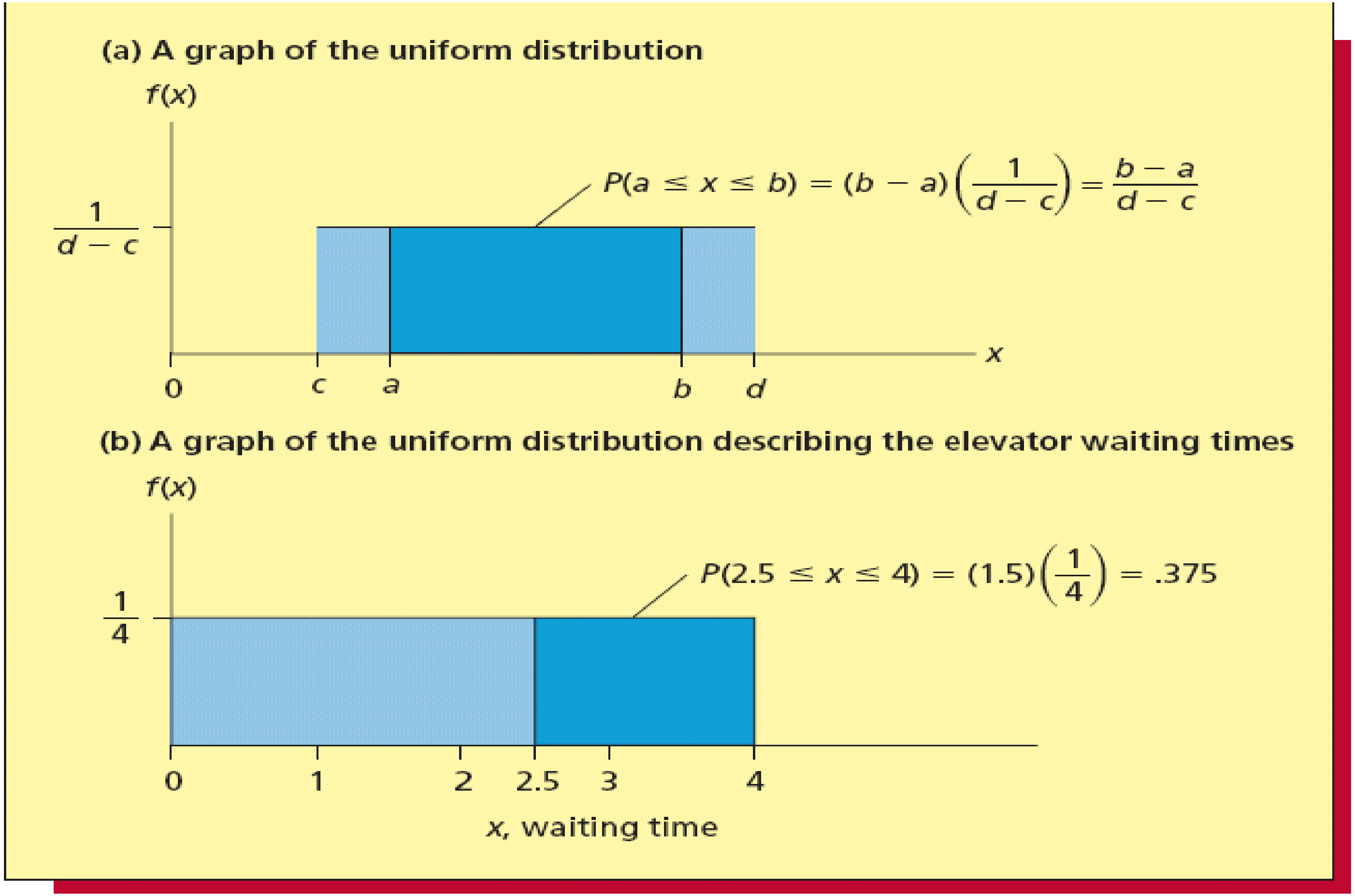


The probability that x is any value between the given values a and b ($a < b$) is

$$P(a \le x \le b) = \frac{b-a}{d-c}$$

Note: The number ordering is $c < a < b < d$

The Uniform Probability Curve



The Uniform Distribution

The mean μ_X and standard deviation σ_X of a uniform random variable *x* are

$$\mu_X = \frac{c+d}{2}$$

$$\sigma_X = \frac{d-c}{\sqrt{12}}$$

Notes on the Uniform Distribution

- The uniform distribution is symmetrical
  - Symmetrical about its center μ_X
  - μ_X is the median

- The uniform distribution is rectangular
  - For endpoints c and d ($c < d$ and $c \neq d$) the width of the distribution is $d - c$ and the height is $1/(d - c)$
  - The area under the entire uniform distribution is 1
    - Because width × height = $(d - c) \times [1/(d - c)] = 1$
    - So $P(c \leq x \leq d) = 1$

Example 6.1 Uniform Waiting Time #1

- The amount of time, x, that a randomly selected hotel patron spends waiting for the elevator at dinnertime.
- Record suggests x is uniformly distributed between zero and four minutes
  - So $c = 0$ and $d = 4$,

$$f(x) = \begin{cases} \dfrac{1}{4-0} = \dfrac{1}{4} & \text{for } 0 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

$$P(a \le x \le b) = \frac{b-a}{4-0} = \frac{b-a}{4}$$

Uniform Waiting Time #2

- What is the probability that a randomly selected hotel patron waits at least 2.5 minutes for the elevator?
  - The probability is the area under the uniform distribution in the interval [2.5, 4] minutes
    - The probability is the area of a rectangle with height ¼ and base 4 – 2.5 = 1.5
  - $P(x \geq 2.5) = P(2.5 \leq x \leq 4) = \frac{1}{4} \times 1.5 = 0.375$

- What is the probability that a randomly selected hotel patron waits less than one minutes for the elevator?

Uniform Waiting Time #3

- Expect to wait the mean time μ_X

$$\mu_X = \frac{0+4}{2} = 2 \text{ minutes}$$

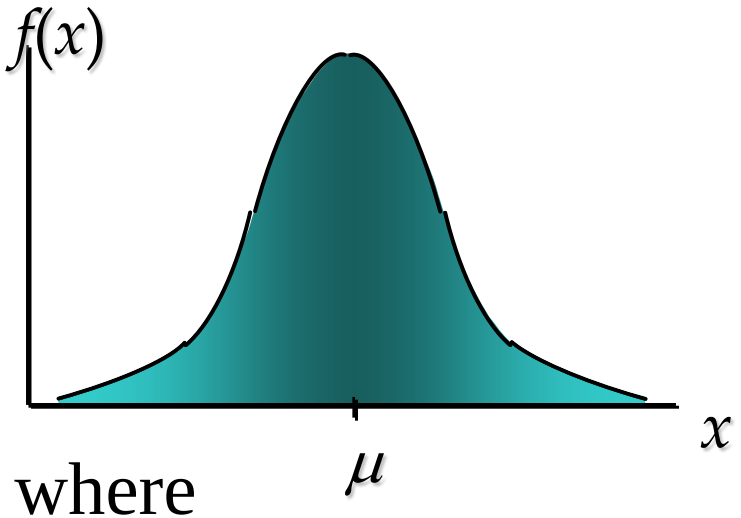
- with a standard deviation σ_X of

$$\sigma_X = \frac{4-0}{\sqrt{12}} = 1.1547 \text{ minutes}$$

6.3 The Normal Distribution

The normal probability distribution is defined by the equation

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



for all values x on the real number line, where

μ is the mean and σ is the standard deviation,
$\pi = 3.14159$,
$e = 2.71828$ is the base of natural logarithms

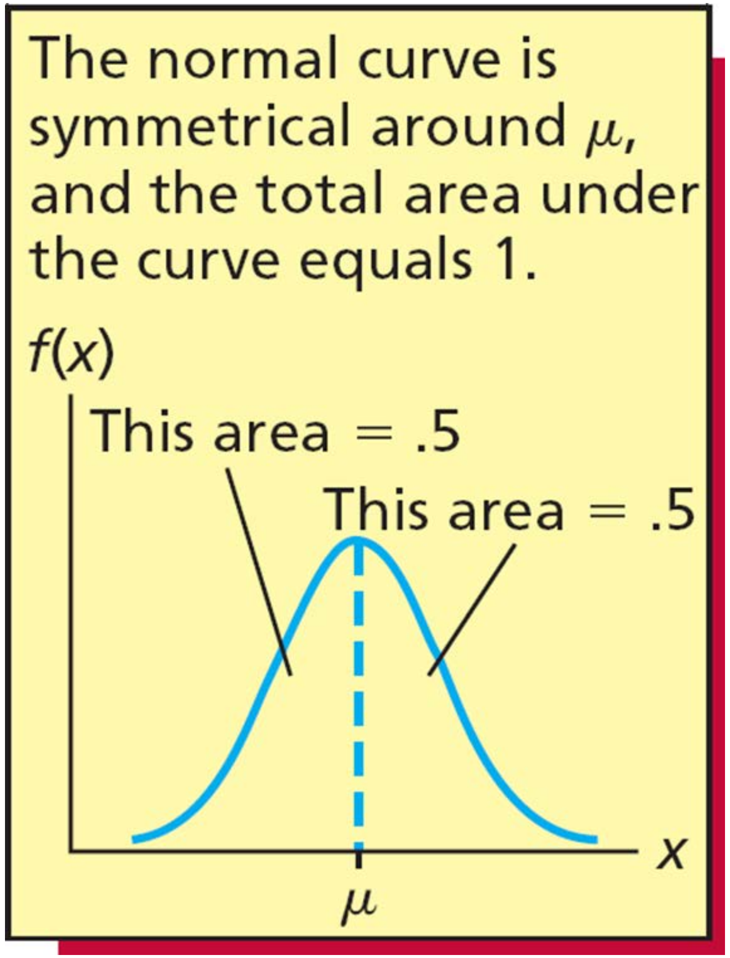

The normal curve is symmetrical around μ, and the total area under the curve equals 1.
$f(x)$
This area = .5
This area = .5
x
μ

Figure 6.3

Properties of the Normal Distribution

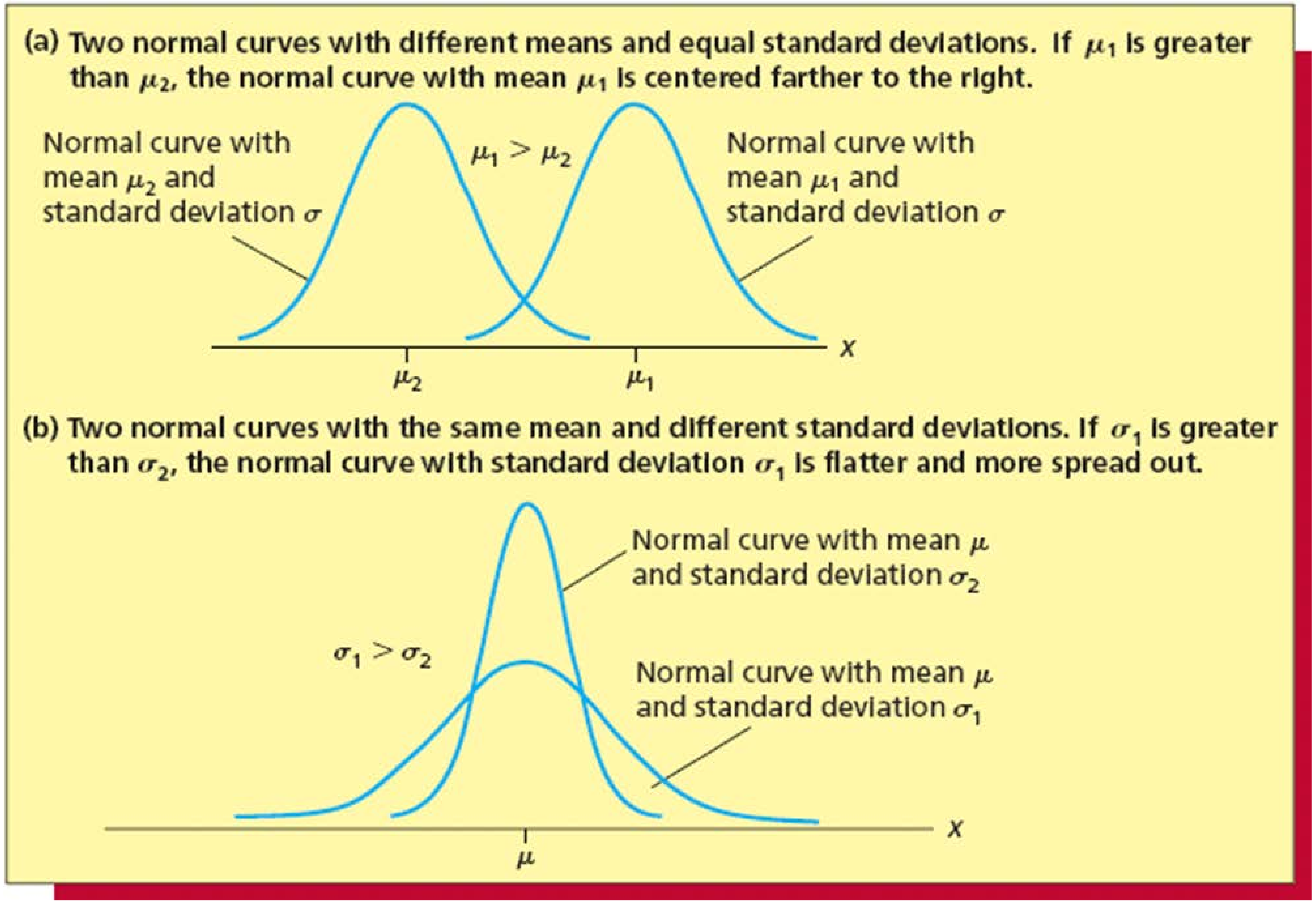
1. There are an infinite number of normal curves
   - The shape of any individual normal curve depends on its specific mean and standard deviation
2. The highest point is over the mean
   - Also the median and mode

Properties of the Normal Distribution

3. The curve is symmetrical about its mean
    - The left and right halves of the curve are mirror images of each other
4. The tails of the normal extend to infinity in both directions
    - The tails get closer to the horizontal axis but never touch it
5. The area under the normal curve to the right of the mean equals the area under the normal to the left of the mean
    - The area under each half is 0.5

The Position and Shape of the Normal Curve

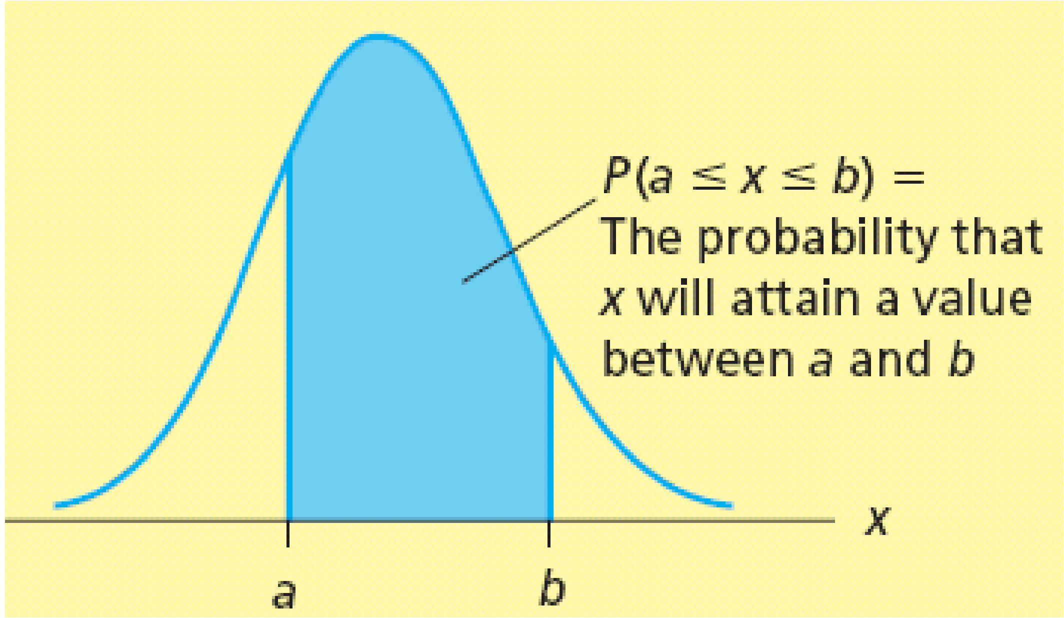
Figure 6.4



Normal Probabilities

Suppose x is a normally distributed random variable with mean μ and standard deviation σ

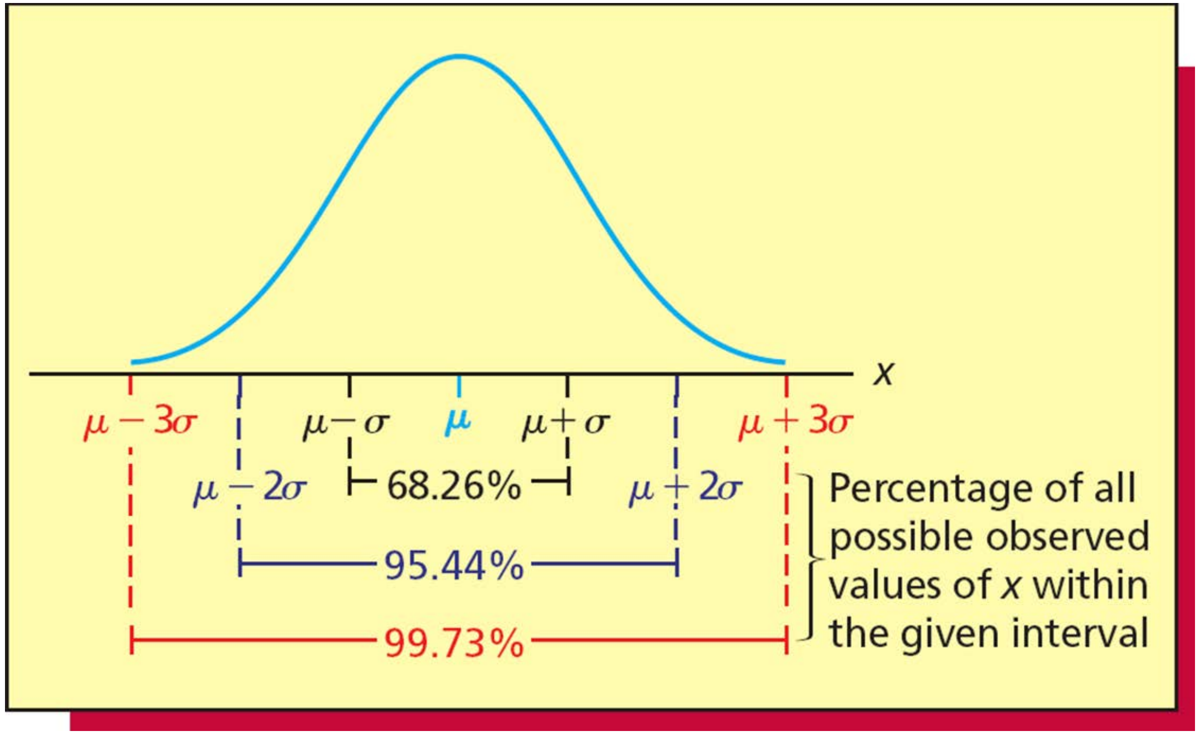
The probability that x could take any value in the range between two given values a and b ($a < b$) is $P(a \leq x \leq b)$



$P(a \leq x \leq b)$ is the area colored in blue under the normal curve and between the values $x = a$ and $x = b$

Three Important Percentages

Figure 6.6



Three Important Areas under the Normal Curve

1. $P(\mu - \sigma \leq x \leq \mu + \sigma) = .6826$

   This means that 68.26 percent of all possible observed values of x are within (plus or minus) one standard deviation of μ.

2. $P(\mu - 2\sigma \leq x \leq \mu + 2\sigma) = .9544$

   This means that 95.44 percent of all possible observed values of x are within (plus or minus) two standard deviations of μ.

3. $P(\mu - 3\sigma \leq x \leq \mu + 3\sigma) = .9973$

   This means that 99.73 percent of all possible observed values of x are within (plus or minus) three standard deviations of μ.

z-scores

- For any *x* in a population or sample, the ***z*-score** corresponding to *x* is defined as

$$z = \frac{x - \text{mean}}{\text{standard deviation}}$$

- The ***z*-score**, which is also called the *standardized value*, is the number of standard deviations that *x* is from the mean
  - A positive *z*-score is for *x* above the mean
  - A negative *z*-score is for *x* below the mean
  - The mean has a *z*-score of zero

The Standard Normal Distribution

If x is normally distributed with mean μ and standard deviation σ, then the random variable z

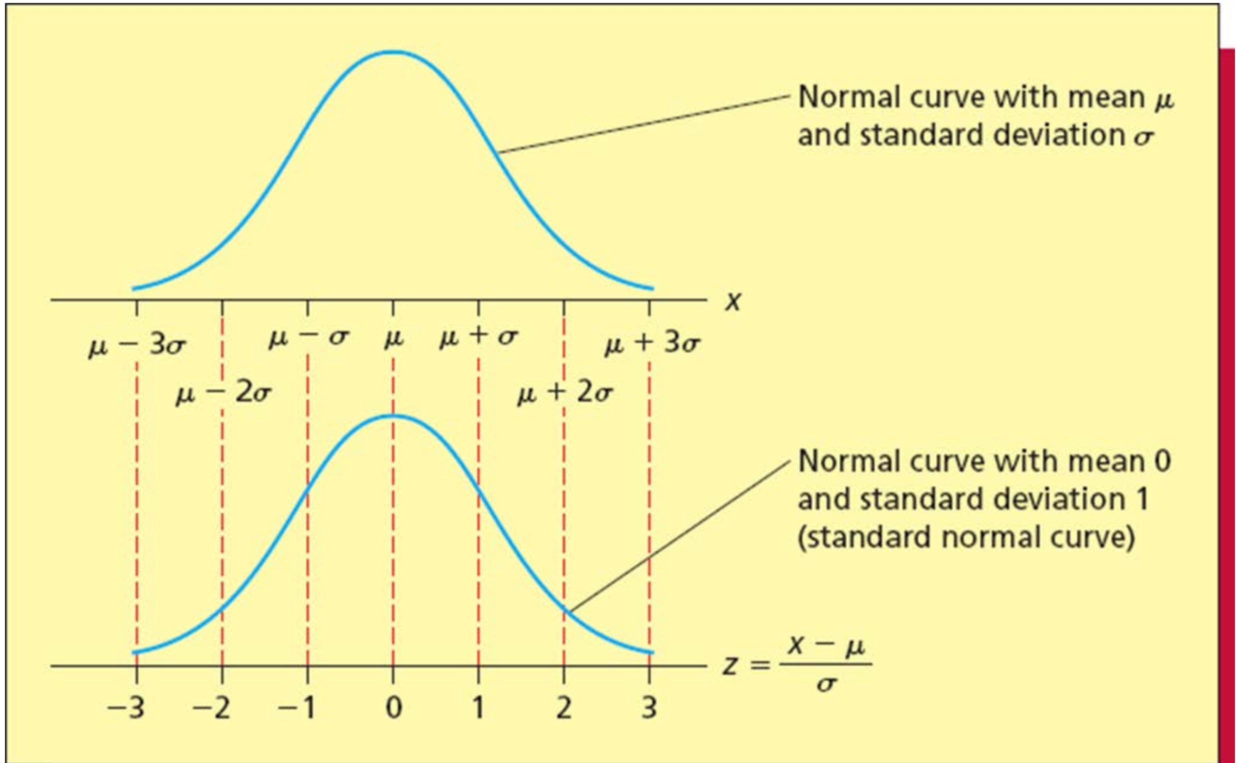
$$z = \frac{x - \mu}{\sigma}$$

is normally distributed with mean 0 and standard deviation 1; this normal is called the standard normal distribution.

The Standard Normal Distribution

z measures the number of standard deviations that x is from the mean μ

- The algebraic sign on z indicates on which side of μ is with x
- z is positive if $x > \mu$ (x is to the right of μ on the number line)
- z is negative if $x < \mu$ (x is to the left of μ on the number line)



The Standard Normal Table

- The standard normal table is a table that lists the cumulative areas under the standard normal curve.
  - This table is very important.
  - Always look at the accompanying figure for guidance on how to use the table

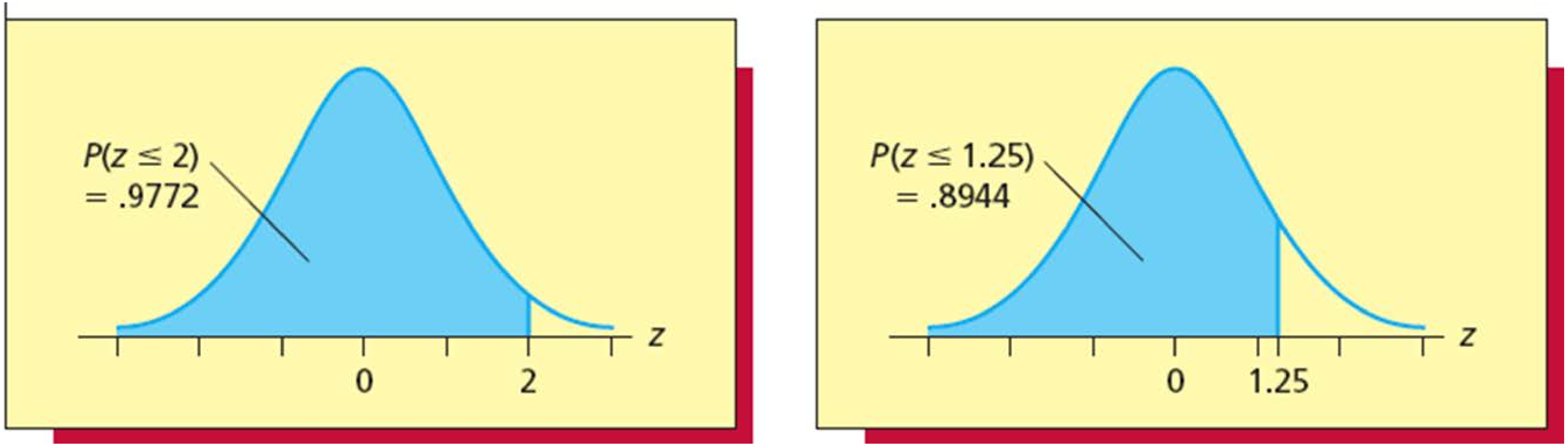
The Cumulative Normal Table

Top of Table 6.1

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−3.9	0.00005	0.00005	0.00004	0.00004	0.00004	0.00004	0.00004	0.00004	0.00003	0.00003
−3.8	0.00007	0.00007	0.00007	0.00006	0.00006	0.00006	0.00006	0.00005	0.00005	0.00005
−3.7	0.00011	0.00010	0.00010	0.00010	0.00009	0.00009	0.00008	0.00008	0.00008	0.00008
−3.6	0.00016	0.00015	0.00015	0.00014	0.00014	0.00013	0.00013	0.00012	0.00012	0.00011
−3.5	0.00023	0.00022	0.00022	0.00021	0.00020	0.00019	0.00019	0.00018	0.00017	0.00017
−3.4	0.00034	0.00032	0.00031	0.00030	0.00029	0.00028	0.00027	0.00026	0.00025	0.00024
−3.3	0.00048	0.00047	0.00045	0.00043	0.00042	0.00040	0.00039	0.00038	0.00036	0.00035
−3.2	0.00069	0.00066	0.00064	0.00062	0.00060	0.00058	0.00056	0.00054	0.00052	0.00050
−3.1	0.00097	0.00094	0.00090	0.00087	0.00084	0.00082	0.00079	0.00076	0.00074	0.00071
−3.0	0.00135	0.00131	0.00126	0.00122	0.00118	0.00114	0.00111	0.00107	0.00103	0.00100
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183

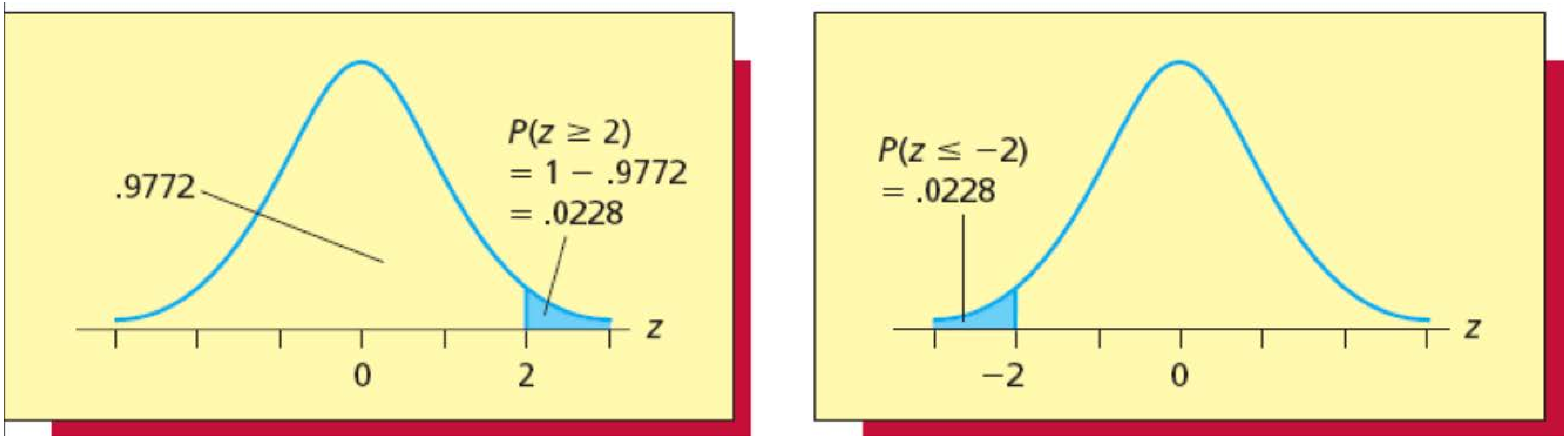
Examples

Figures 6.8 and 6.9



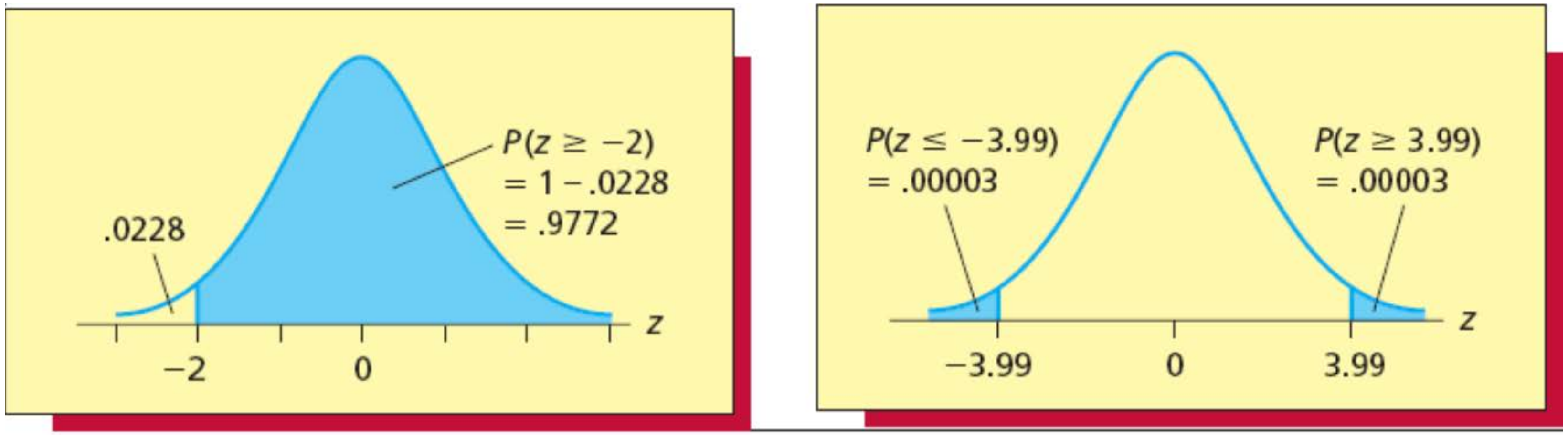
Examples

Figures 6.10 and 6.11



Examples

Figures 6.12 and 6.13

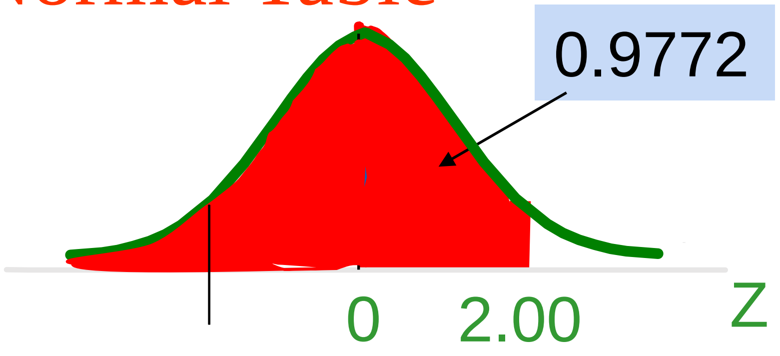


The Standard Normal Table

- The values of z (accurate to the nearest tenth) in the table range from -3.99 to 3.99 in increments of 0.01
- The areas under the normal curve to the left of any value of z are given in the body of the table

The Standardized Normal Table

- The Standardized Normal table in the textbook gives the probability that z will be less than or equal to 2.00



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

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

z	0.00	0.01	0.02 …	0.06
0.0				
0.1				
.				
.				
1.9.				0.9750
2.0	. 9772			

P(Z <= 2.00) = 0.9772, P(Z<=1.96)=0.9750

- Find $P(z \leq 2)$
    - Find the area listed in the table corresponding to a z value of 2.00
    - Starting from the top of the far left column, go down to "2.0"
    - Read across the row $z = 2.0$ until under the column headed by ".00"
    - The area is in the cell that is the intersection of this row with this column
    - As listed in the table, the area is 0.9772, so

$$P(z \leq 2) = 0.9772$$

Calculating $P(-2.53 \leq z \leq 2.53)$

- First, find $P(z \leq 2.53)$
  - Go to the table of areas under the standard normal curve
  - Go down left-most column for $z = 2.5$
  - Go across the row 2.5 to the column headed by .03
  - The cumulative area to the value of $z = 2.53$ is the value contained in the cell that is the intersection of the 2.5 row and the .03 column
  - The table value for the area is 0.9943

Calculating $P\ (-2.53 \le z \le 2.53)$

- From last slide, $P\ (z \le 2.53)=0.9943$
- By symmetry of the normal curve,

  $P(z \le -2.53)=1-0.9943=0.0057$
- Then $P\ (-2.53 \le z \le 2.53)$
- $= P\ (z \le 2.53)-P(z \le -2.53)$

  $= 0.9943 - 0.0057 = 0.9886$

Alternative:

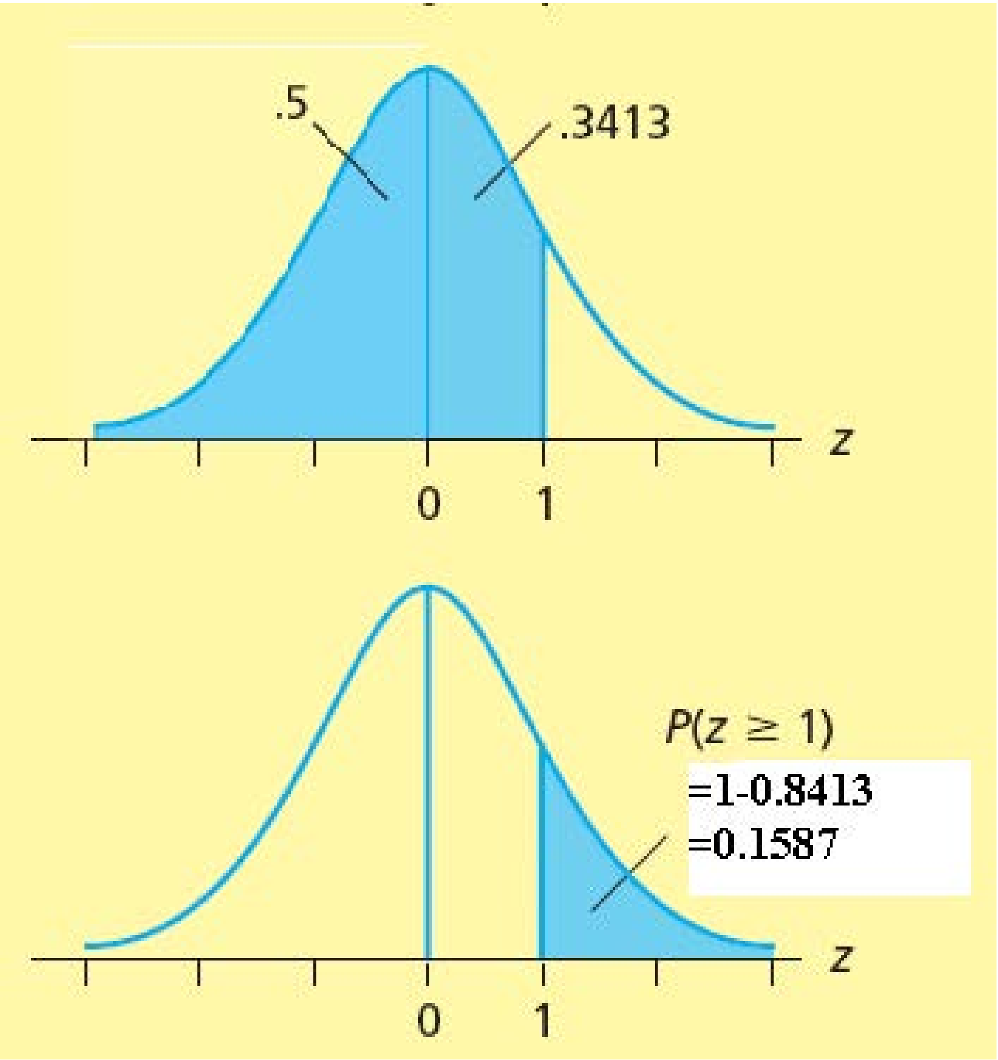
$P\ (0 \le z \le 2.53)=0.9943-0.5=0.4943$

$P\ (-2.53 \le z \le 2.53)=0.4943+0.4943=0.9886$

Calculating $P(z \geq 1)$

An example of finding tail areas

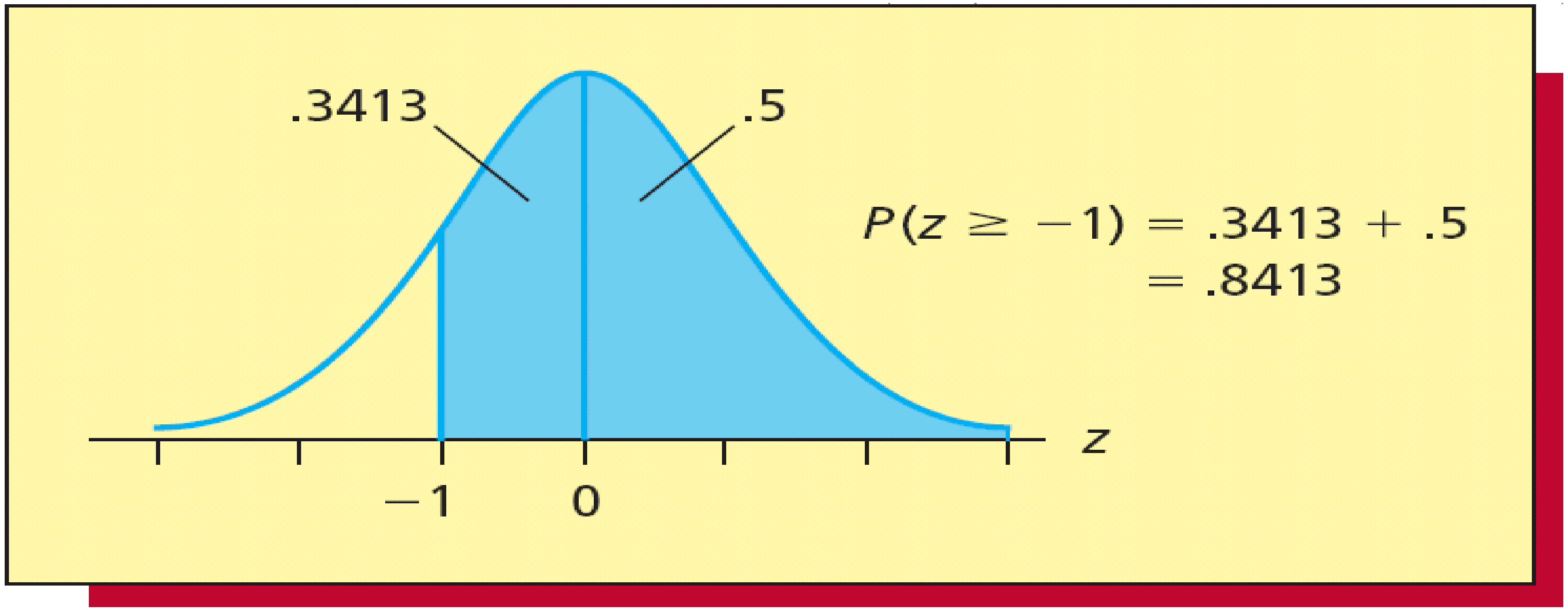
- Shown is finding the right-hand tail area for $z \geq 1.00$
  - Equivalent to the left-hand tail area for $z \leq -1.00$



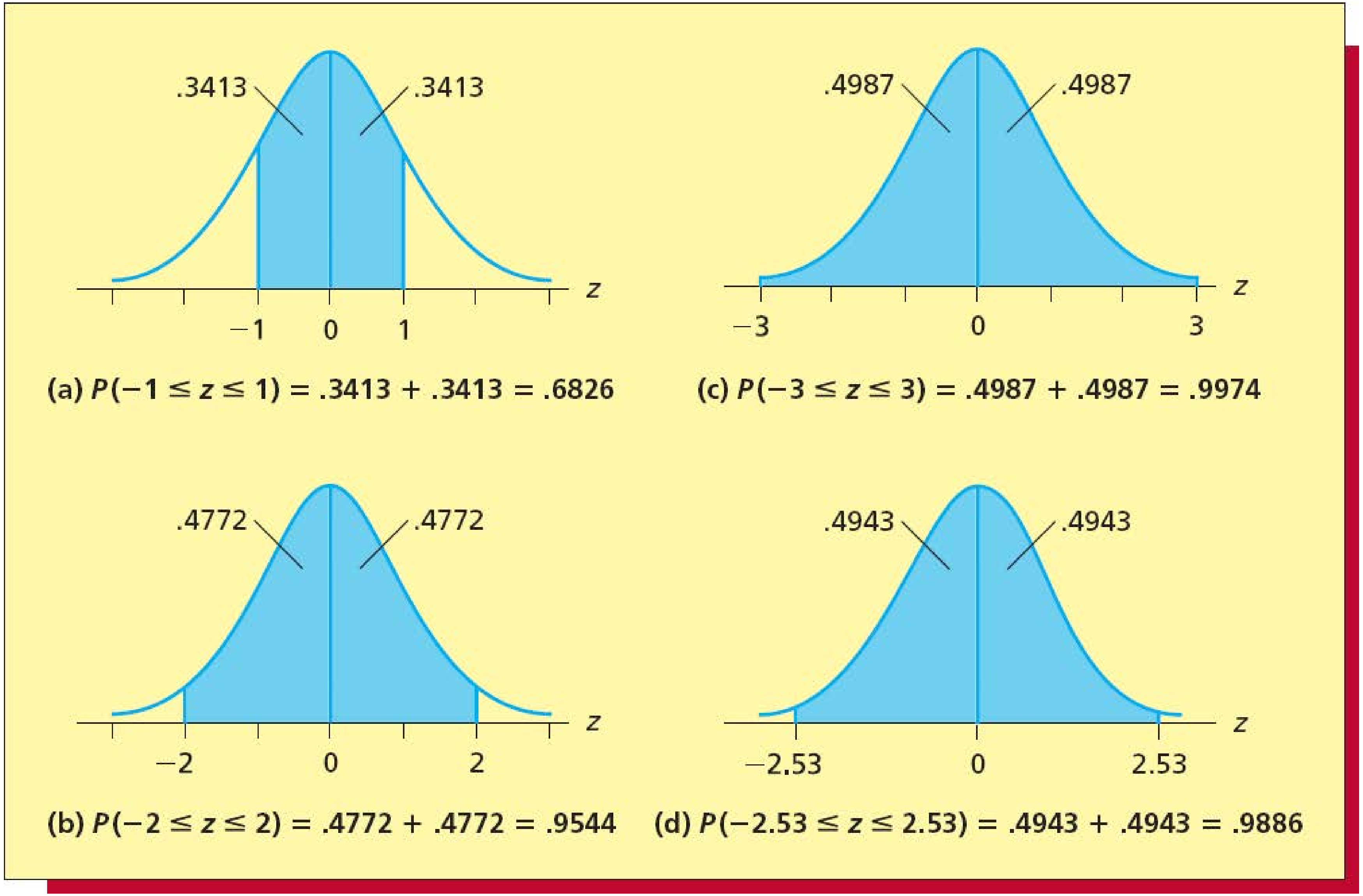
Calculating $P\ (z \geq -1)$

An example of finding the area under the standard normal curve to the right of a negative z value

- Shown is finding the under the standard normal for $z \geq -1$



Some Areas under the Standard Normal Curve



Finding Normal Probabilities

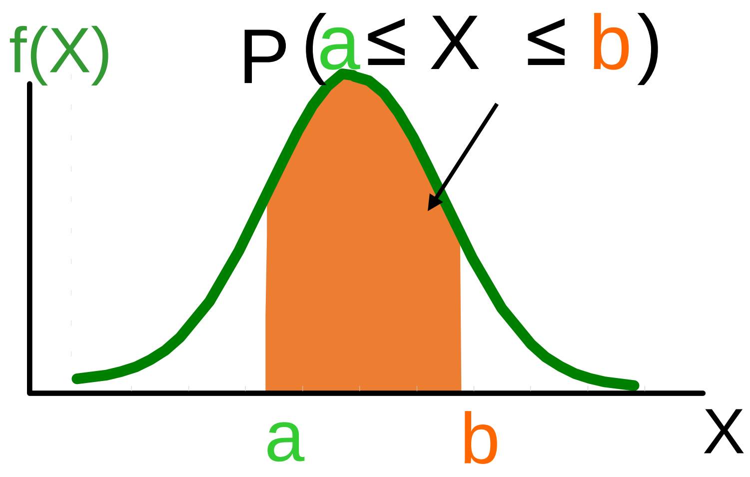
General procedure:

1. Formulate the problem in terms of x values
2. Calculate the corresponding z values, and restate the problem in terms of these z values
3. Find the required areas under the standard normal curve by using the table

   Note: It is always useful to draw a picture showing the required areas before using the normal table

Finding Normal Probabilities

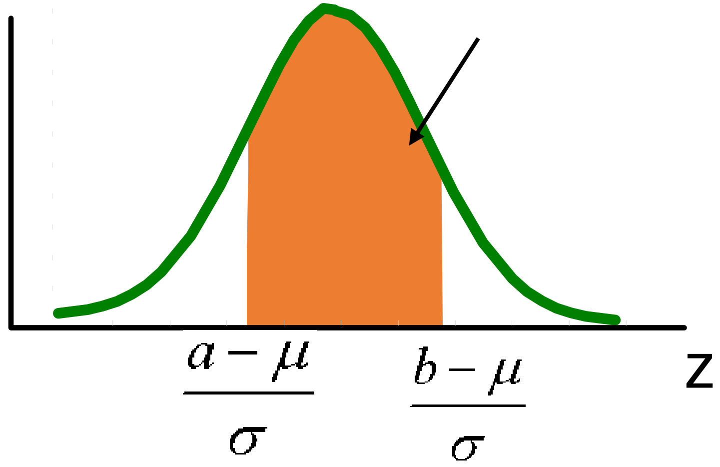
Probability is the area under the curve!



X is normal random variable with mean μ and s.d. σ

$$Z = \frac{X - \mu}{\sigma}$$

Z follows standard normal distribution



$$P(a \le X \le b)$$

$$= P\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right)$$

Three Important Areas under the Normal Curve

1. $P(\mu - \sigma \leqslant x \leqslant \mu + \sigma) = 0.6826$
   - So 68.26% of all possible observed values of x are within (plus or minus) one standard deviation of μ
2. $P(\mu - 2\sigma \leqslant x \leqslant \mu + 2\sigma) = 0.9544$
   - So 95.44% of all possible observed values of x are within (plus or minus) two standard deviations of μ
3. $P(\mu - 3\sigma \leqslant x \leqslant \mu + 3\sigma) = 0.9973$
   - So 99.73% of all possible observed values of x are within (plus or minus) three standard deviations of μ

The Empirical Rule for Normal Populations

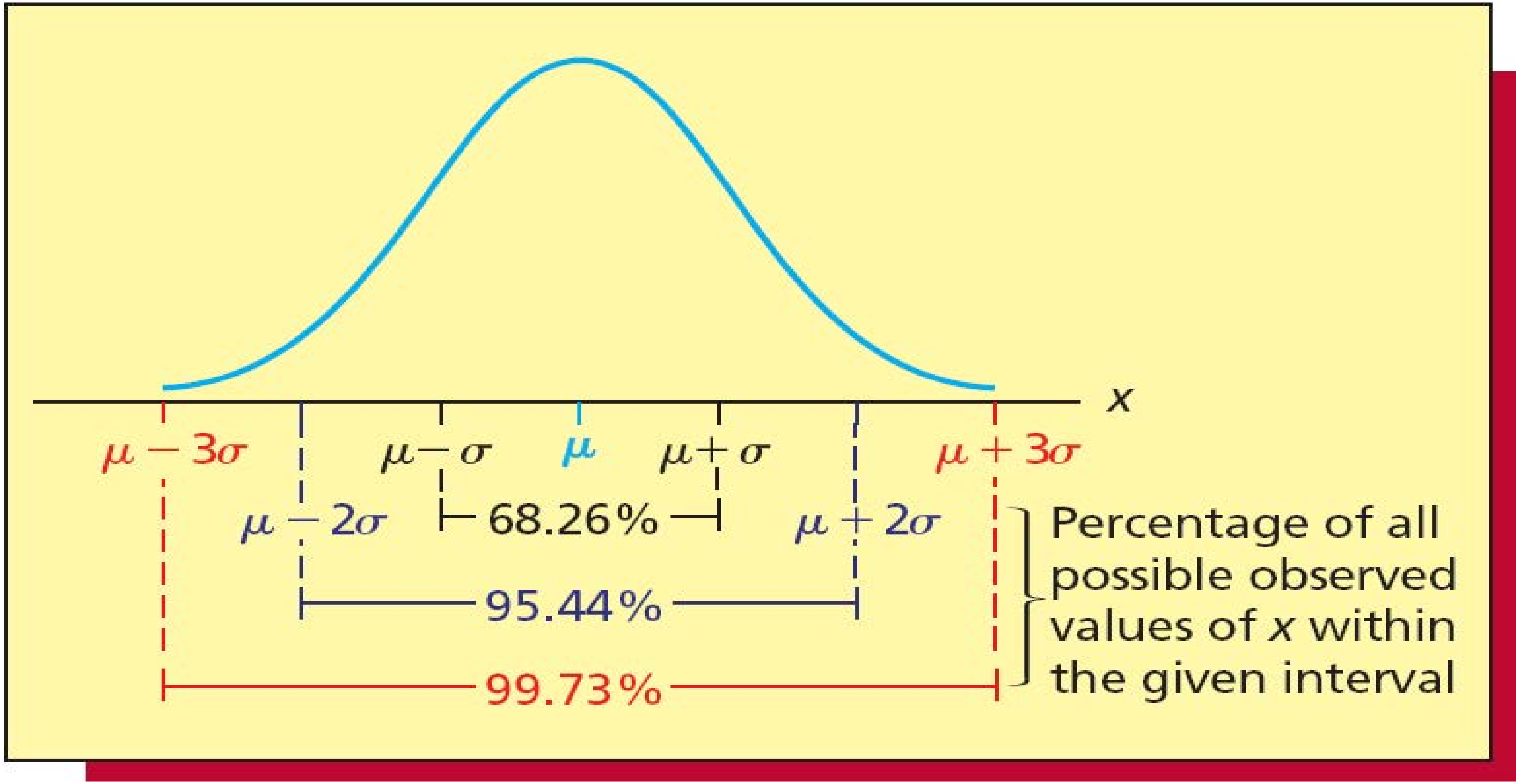
If a population has mean μ and standard deviation σ and is described by a normal curve (symmetrical, bell-shaped distribution), then,

1. 68.26% of the population measurements lie within one standard deviation of the mean: $[\mu-\sigma, \mu+\sigma]$

2. 95.44% of the population measurements lie within two standard deviations of the mean: $[\mu-2\sigma, \mu+2\sigma]$

3. 99.73% of the population measurements lie within three standard deviations of the mean: $[\mu-3\sigma, \mu+3\sigma]$

Three Important Areas under the Normal Curve
Empirical Rule (Visually)



Example 6.2 The Car Mileage Case

Want the probability that the mileage of a randomly selected midsize car will be between 32 and 35 mpg

- Let x be the random variable of mileage of midsize cars, in mpg
- Want $P\ (32 \leq x \leq 35$ mpg)

Given: x is normally distributed

$\mu = 33$ mpg

$\sigma = 0.7$ mpg

The Car Mileage Case #2

Procedure (from previous slide):

1. Formulate the problem in terms of x (as above)

2. Restate in terms of corresponding z values (see next slide)

3. Find the indicated area under the standard normal curve using the normal table (see the slide after that)

The Car Mileage Case #3

For x = 32 mpg, the corresponding z value is

$$z = \frac{x-\mu}{\sigma} = \frac{32-33}{0.7} = \frac{-1}{0.7} = -1.43$$

(so the mileage of 32 mpg is 1.43 standard deviations below (to the left of) the mean μ = 32 mpg)

For x = 35 mpg, the corresponding z value is

$$z = \frac{x-\mu}{\sigma} = \frac{35-33}{0.7} = \frac{2}{0.7} = 2.86$$

(so the mileage of 35 mpg is 2.86 standard deviations above (to the right of) the mean μ = 32 mpg)

Then $P(32 \leq x \leq 35 \text{ mpg}) = P(-1.43 \leq z \leq 2.86)$

The Car Mileage Case #4

- Want: the area under the normal curve between 32 and 35 mpg
- Will find: the area under the standard normal curve between -1.43 and 2.86

The Car Mileage Case #5

- Break this into two pieces, around the mean μ
  - The cumulative area to the left of μ between -1.43 and 0
    - By symmetry, this is the same as the area between 0 and 1.43
    - From the standard normal table, this area is

      0.9236-0.5=0.4236
  - The area to the right of μ between 0 and 2.86
    - From the standard normal table, this area is

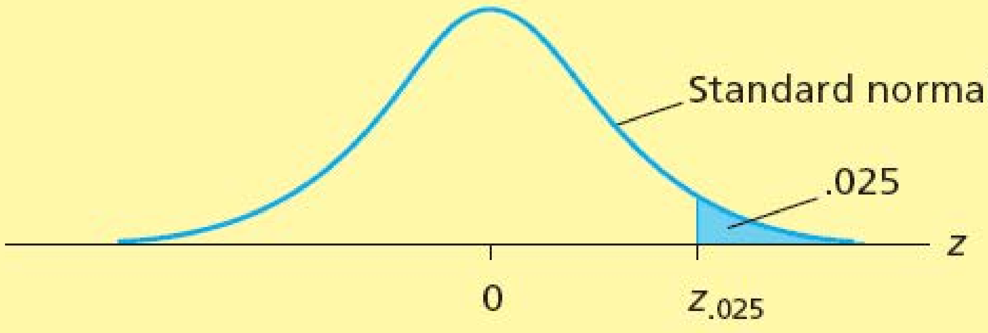
      0.9979-0.5=0.4979
  - The total area of both pieces is 0.4236 + 0.4979 = 0.9215
  - Then, $P(-1.43 \leq z \leq 2.86) = 0.9215$
- Returning to x, $P(32 \leq x \leq 35 \text{ mpg}) = 0.9215$

Exercise

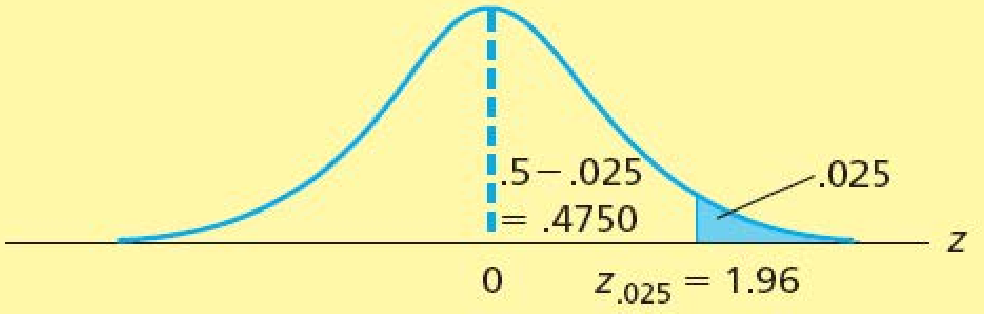
The daily water usage per person in Providence, Rhode Island is normally distributed with a mean of 20 gallons and a standard deviation of 5 gallons. What percent of the population use between 18 and 26 gallons per day?

Solution

Finding *z* Points on a Standard Normal Curve



(a) $z_{.025}$ is the point on the horizontal axis under the standard normal curve that gives a right-hand tail area equal to .025



(b) Finding $z_{.025}$

Example — Stocking out of inventory

A large discount store sells blank VHS tapes want to know how many blank VHS tapes to stock (at the beginning of the week) so that there is only a 5 percent chance of stocking out during the week

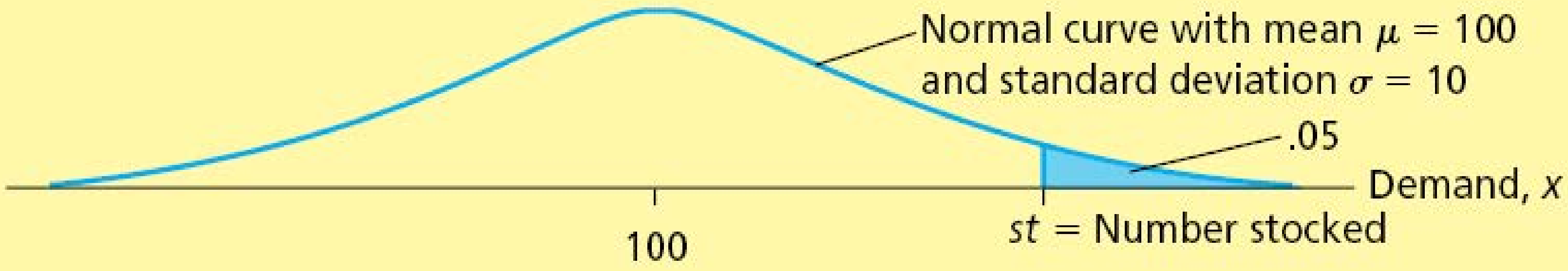
- Let x be the random variable of weekly demand
- Let st be the number of tapes so that there is only a 5% probability that weekly demand will exceed st
- Want the value of st so that $P(x \geq st) = 0.05$

Given: x is normally distributed

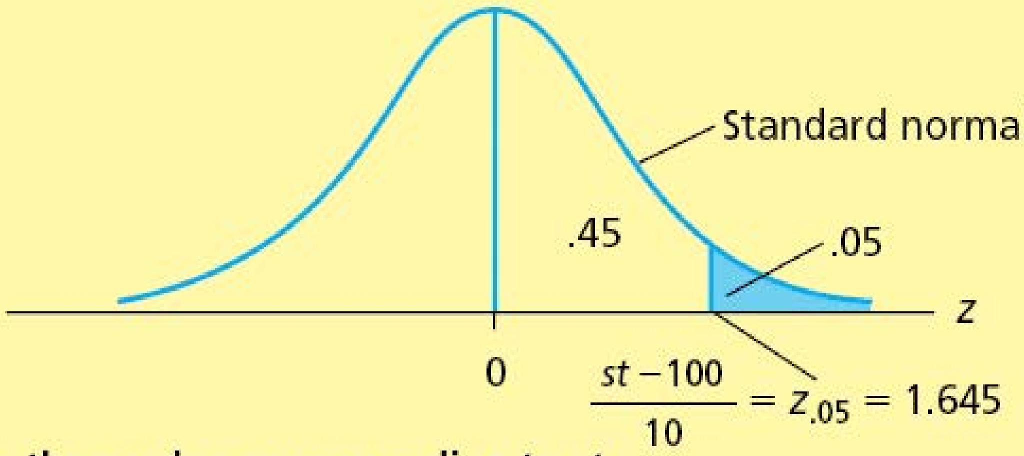
$\mu = 100$ tapes

$\sigma = 10$ tapes

Finding Z Points on a Normal Curve



(a) The number of tapes stocked, st, must be chosen so that there is a .05 probability that the demand, x, will exceed st



(b) Finding $z_{.05}$, the z value corresponding to st

Example #1

- Refer to the above figure.
- In panel (a), *st* is located on the horizontal axis under the right-tail of the normal curve having mean $\mu = 100$ and standard deviation $\sigma = 10$
- The *z* value corresponding to *st* is

$$z = \frac{st - \mu}{\sigma} = \frac{st - 100}{10}$$

- In panel (b), the right-tail area is 0.05, so the cumulative area under the standard normal curve to *z* is 1 – 0.05 = 0.95

Example #2

- Use the standard normal table to find the *z* value associated with a table entry of 0.95
  - But do not find 0.95; instead find values that bracket 0.95
  - For a table entry of 0.9495, *z* = 1.64
  - For a table entry of 0.9505, *z* = 1.65
  - For an area of 0.95, use the *z* value midway between these
  - So *z* = 1.645

$$\frac{st-\mu}{\sigma} = \frac{st-100}{10} = 1.645$$

Solving for *s* gives *st* = 100 + (1.645 × 10) = 116.45
Rounding up, 117 tapes should be stocked so that the probability of running out will not be more than 5 percent

Exercise

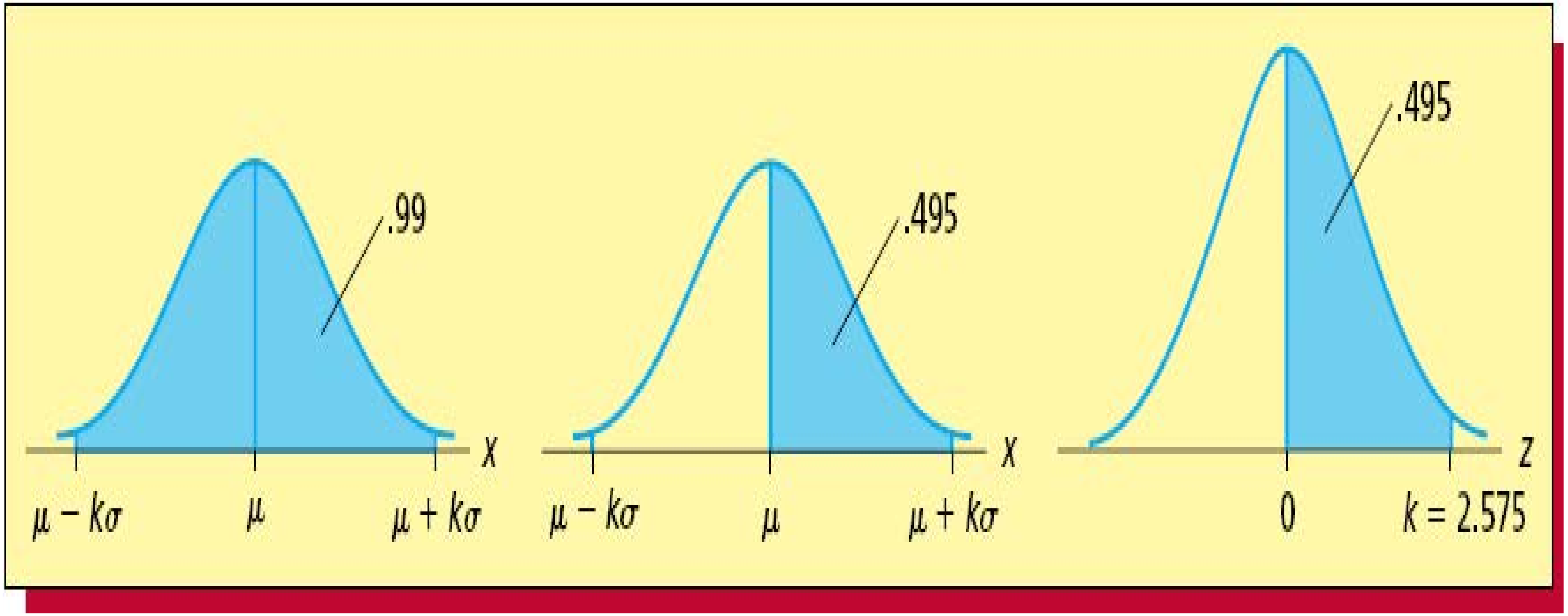
Suppose X is normal with mean 8.0 and standard deviation 5.0

- Find $P(X < 8.6)$
- Find $P(8 < X < 8.6)$
- Find $P(7.4 < X < 8)$
- Find the X value so that only 20% of all values are below this X

Solution:

Finding a Tolerance Interval

Finding a tolerance interval $[\mu \pm k\sigma]$ that contains 99% of the measurements in a normal population



Example 6.4

The Coffee Temperature Case: Meeting Customer Requirements C

Recall that marketing research done by a fast-food restaurant indicates that coffee tastes best if its temperature is between 153°(F) and 167°(F). The restaurant has sampled the coffee it serves and observed the 48 temperature readings in Table 1.10 on page 17. The temperature readings have a mean $\bar{x} = 159.3958$ and a standard deviation $s = 6.4238$ and are described by a bell-shaped histogram. Using $\bar{x}$ and s as point estimates of the mean μ and the standard deviation σ of the population of all possible coffee temperatures, we wish to calculate the probability that x, the temperature of a randomly selected cup of coffee, is outside the customer requirements for best-testing coffee (that is, less than 153° or greater than 167°). In order to compute the probability $P(x < 153 \text{ or } x > 167)$, we compute the z values

$$z = \frac{153 - 159.3958}{6.4238} = -1.00 \qquad \text{and} \qquad z = \frac{167 - 159.3958}{6.4238} = 1.18$$

Because the events $\{x < 153\}$ and $\{x > 167\}$ are mutually exclusive, we have

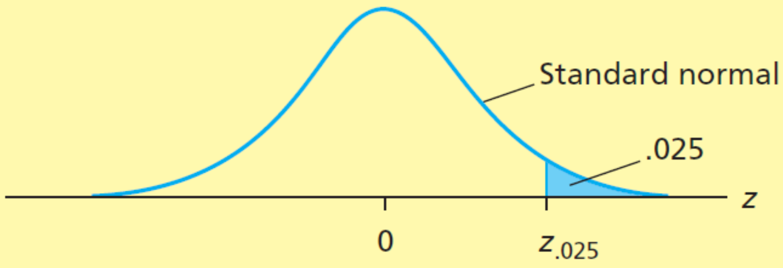
$$\begin{aligned} P(x < 153 \text{ or } x > 167) &= P(x < 153) + P(x > 167) \\ &= P(z < -1.00) + P(z > 1.18) \\ &= .1587 + .1190 = .2777 \end{aligned}$$

BI

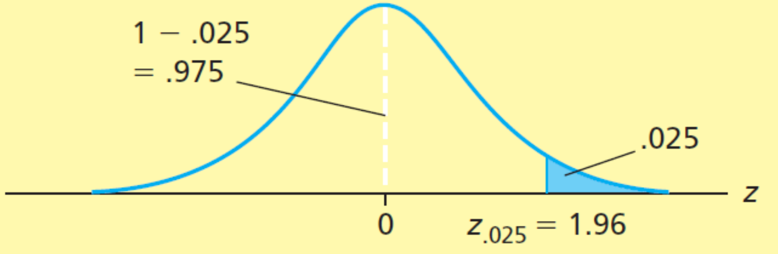
This calculation is illustrated in Figure 6.18. The probability of .2777 implies that 27.77 percent of the coffee temperatures do not meet customer requirements and 72.23 percent of the coffee temperatures do meet these requirements. If management wishes a very high percentage of its coffee temperatures to meet customer requirements, the coffee-making process must be improved.

In general, **we let z_α denote the point on the horizontal axis under the standard normal curve that gives a right-hand tail area equal to α.** With this definition in mind, we consider the following example.

FIGURE **The Point $z_{.025} = 1.96$**



(a) $z_{.025}$ is the point on the horizontal axis under the standard normal curve that gives a right-hand tail area equal to .025



(b) Finding $z_{.025}$

EXAMPLE The DVD Case: Managing Inventory

A large discount store sells 50 packs of HX-150 blank DVDs and receives a shipment every Monday. Historical sales records indicate that the weekly demand, x, for these 50 packs is normally distributed with a mean of $\mu = 100$ and a standard deviation of $\sigma = 10$. How many 50 packs should be stocked at the beginning of a week so that there is only a 5 percent chance that the store will run short during the week?

If we let st equal the number of 50 packs that will be stocked, then st must be chosen to allow only a .05 probability that weekly demand, x, will exceed st. That is, st must be chosen so that

$$P(x > st) = .05$$

Figure 6.20(a) on the next page shows that the number stocked, st, is located under the right-hand tail of the normal curve having mean $\mu = 100$ and standard deviation $\sigma = 10$. In order to find st, we need to determine how many standard deviations st must be above the mean in order to give a right-hand tail area that is equal to .05.

The z value corresponding to st is

$$z = \frac{st - \mu}{\sigma} = \frac{st - 100}{10}$$

This last equation says that st is 1.645 standard deviations ($\sigma = 10$) above the mean ($\mu = 100$). Rounding $st = 116.45$ up so that the store's chances of running short will be *no more* than 5 percent, the store should stock 117 of the 50 packs at the beginning of each week.

Sometimes we need to find the point on the horizontal axis under the standard normal curve that gives a particular **left-hand tail area** (say, for instance, an area of .025). Looking at Figure 6.21, it is easy to see that, if, for instance, we want a left-hand tail area of .025, the needed z value is $-z_{.025}$, where $z_{.025}$ gives a right-hand tail area equal to .025. To find $-z_{.025}$, we look for .025 in the body of the normal table and find that the z value corresponding to .025 is -1.96. Therefore, $-z_{.025} = -1.96$. In general, **$-z_\alpha$ is the point on the horizontal axis under the standard normal curve that gives a left-hand tail area equal to α.**

EXAMPLE 6.6 Setting A Guarantee Period

Extensive testing indicates that the lifetime of the Everlast automobile battery is normally distributed with a mean of $\mu = 60$ months and a standard deviation of $\sigma = 6$ months. The Everlast's manufacturer has decided to offer a free replacement battery to any purchaser whose Everlast battery does not last at least as long as the minimum lifetime specified in its guarantee. How can the manufacturer establish the guarantee period so that only 1 percent of the batteries will need to be replaced free of charge?

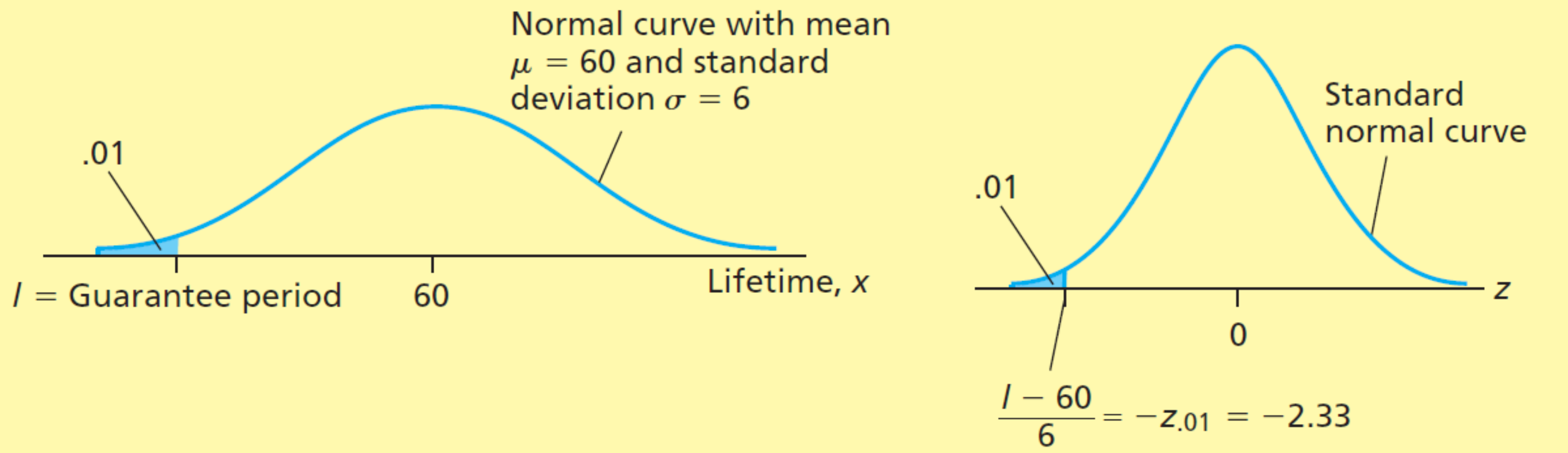
If the battery will be guaranteed to last l months, l must be chosen to allow only a .01 probability that the lifetime, x, of an Everlast battery will be less than l. That is, we must choose l so that

$$P(x < l) = .01$$

Figure 6.22(a) shows that the guarantee period, l, is located under the left-hand tail of the normal curve having mean $\mu = 60$ and standard deviation $\sigma = 6$. In order to find l, we need to determine how many standard deviations l must be below the mean in order to give a left-hand tail area that equals .01. The z value corresponding to l is

$$z = \frac{l - \mu}{\sigma} = \frac{l - 60}{6}$$

FIGURE 6.22 Finding the Guarantee Period, l, so That $P(x < l) = .01$ When $\mu = 60$ and $\sigma = 6$



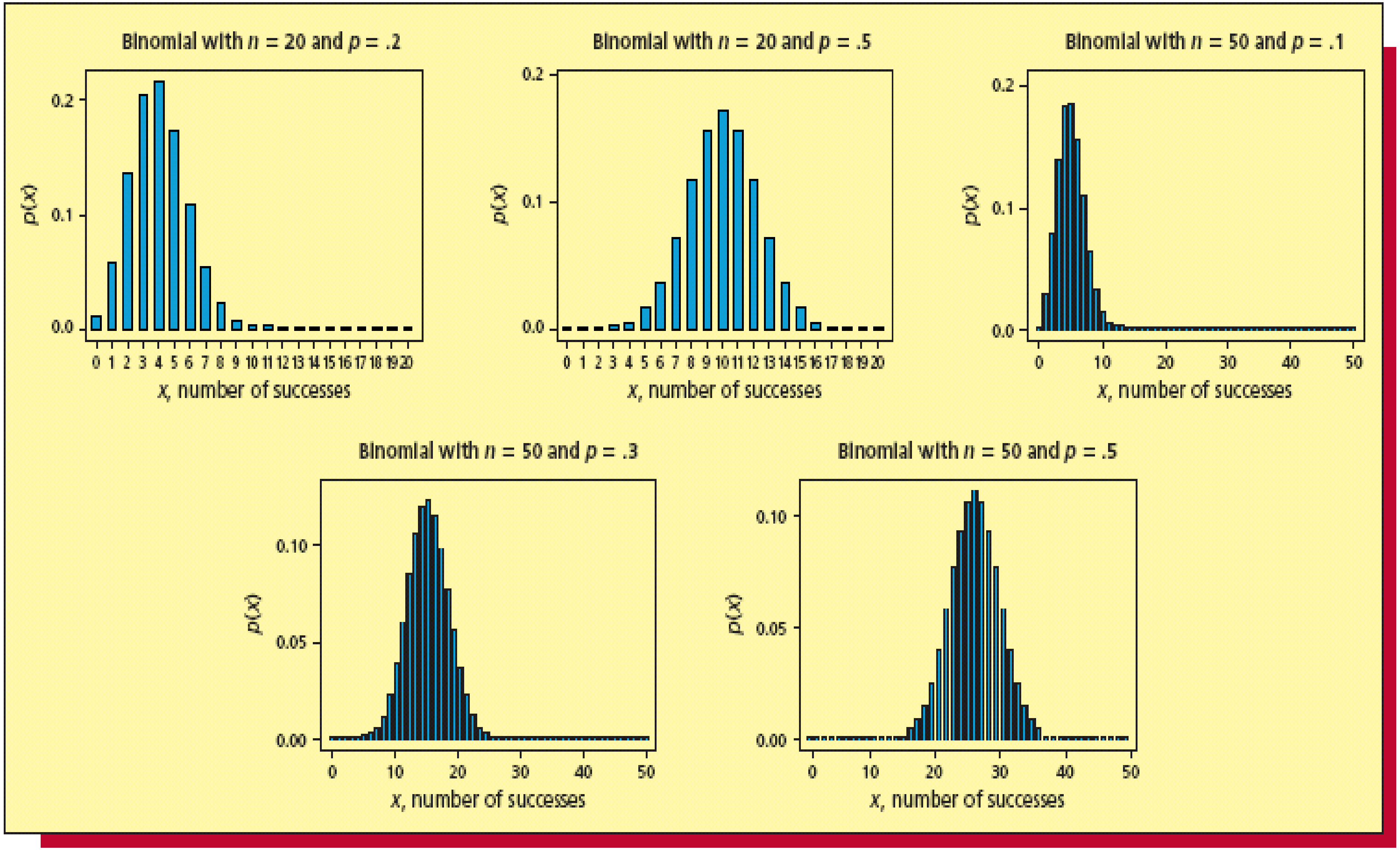
(a) The guarantee period, l, must be chosen so that there is a .01 probability that the lifetime, x, will be less than l

(b) Finding $-z_{.01}$, the z value corresponding to l

6.4 Approximating the Binomial Distribution by Using the Normal Distribution (Optimal)

- The figure below shows several binomial distributions
- Can see that as *n* gets larger and as *p* gets closer to 0.5, the graph of the binomial distribution tends to have the symmetrical, bell-shaped, form of the normal curve

Normal Approximation to the Binomial



Normal Approximation to the Binomial

- Generalize observation from last slide for large n
- Suppose x is a binomial random variable, where n is the number of trials, each having a probability of success p
  - Then the probability of failure is $1 - p$
- If n and p are such that $np \geq 5$ and $n(1 - p) \geq 5$, then x is approximately normal with

$$\mu = np \ \text{ and } \ \sigma = \sqrt{np(1 - p)}$$

Example 6.8

Normal Approximation to a Binomial #1

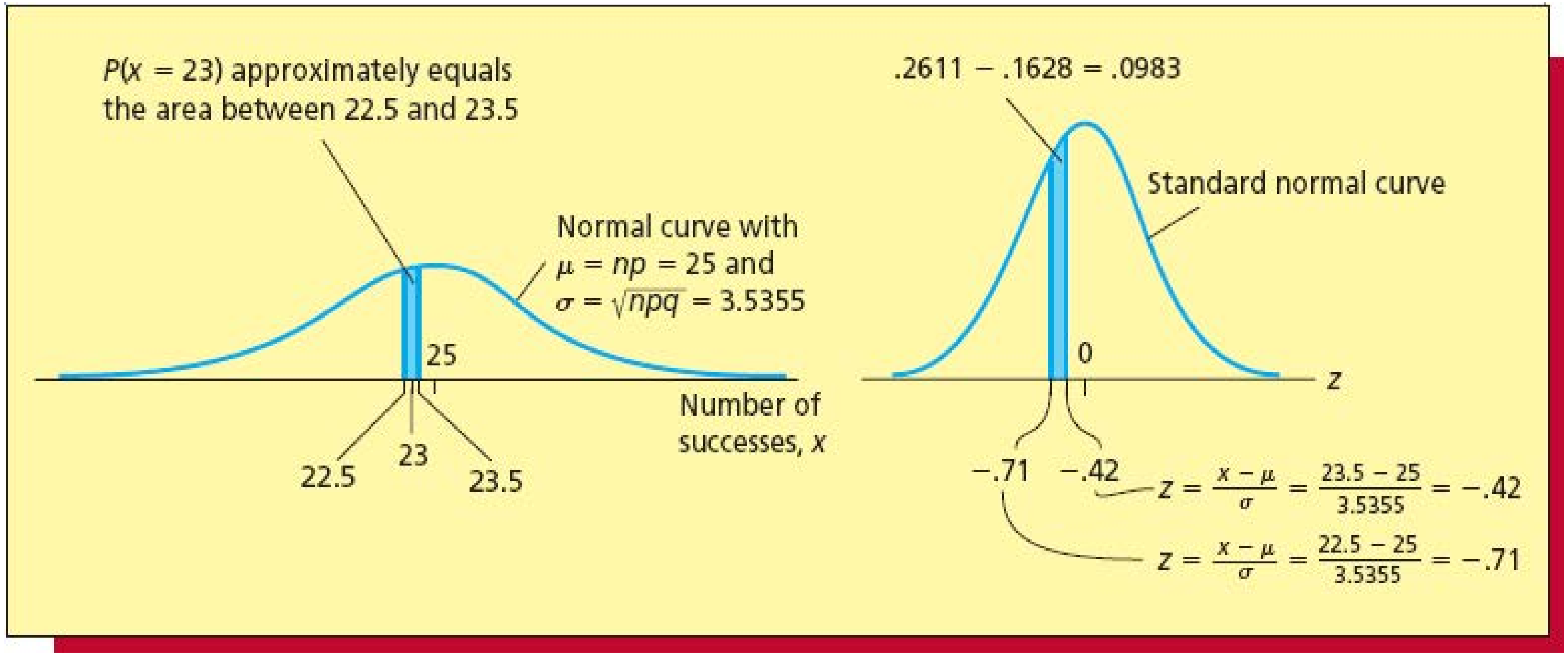
- Suppose there is a binomial variable x with n = 50 and p = 0.5
- Want the probability of x = 23 successes in the n = 50 trials
  - Want $P(x = 23)$
  - Approximating by using the normal curve with

$$\mu = np = (50 \times 0.50) = 25$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{50 \times 0.50 \times 0.50} = 3.5355$$

Normal Approximation to a Binomial #2

- With continuity correction, find the normal probability $P(22.5 \leq x \leq 23.5)$



Normal Approximation to a Binomial #3

- For x = 22.5, the corresponding z value is

$$z = \frac{x - \mu}{\sigma} = \frac{22.5 - 25}{3.5355} = -0.71$$

- For x = 23.5, the corresponding z value is

$$z = \frac{x - \mu}{\sigma} = \frac{23.5 - 25}{3.5355} = -0.42$$

- Then $P(22.5 \le x \le 23.5) = P(-0.71 \le z \le -0.42)$

$= 0.7611 - 0.6628$

$= 0.0983$

- Therefore the estimate of the binomial probability $P(x = 23)$ is 0.0983

1. Professor Mann has determined that the scores in his statistics course are approximately normally distributed with a mean of 72 and a standard deviation of 5. He announces to the class that the top 15 percent of the scores will earn an A. What is the lowest score a student can earn and still receive an A?

2. Consider the binomial random variable x with n = 50 trials and probability of success p = 0.4. Use the normal approximation to this binomial distribution to compute the probability of at least 21 successes in the 50 trails.

TABLE 6.2 Several Examples of the Continuity Correction ($n = 50$)

Binomial Probability	Numbers of Successes Included in Event	Normal Curve Area (with Continuity Correction)
$P(25 < x \leq 30)$	26, 27, 28, 29, 30	$P(25.5 \leq x \leq 30.5)$
$P(x \leq 27)$	0, 1, 2, . . . , 26, 27	$P(x \leq 27.5)$
$P(x > 30)$	31, 32, 33, . . . , 50	$P(x \geq 30.5)$
$P(27 < x < 31)$	28, 29, 30	$P(27.5 \leq x \leq 30.5)$

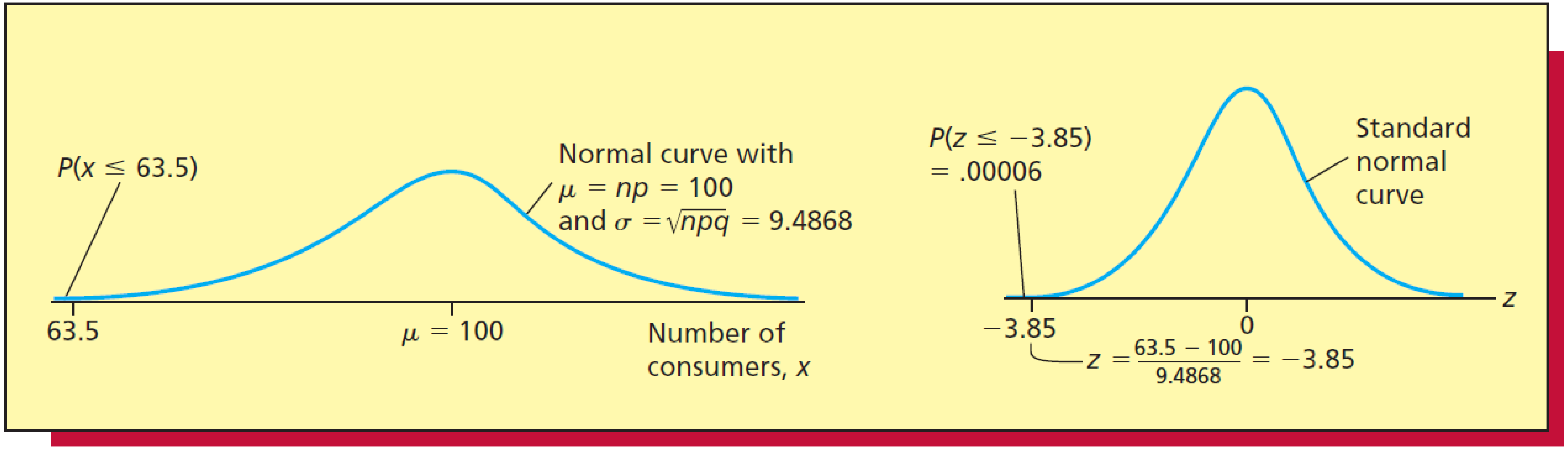
EXAMPLE 6.9 The Cheese Spread Case: Improving Profitability

C

A food processing company markets a soft cheese spread that is sold in a plastic container with an "easy pour" spout. Although this spout works extremely well and is popular with consumers, it is expensive to produce. Because of the spout's high cost, the company has developed a new, less expensive spout. While the new, cheaper spout may alienate some purchasers, a company study shows that its introduction will increase profits if fewer than 10 percent of the cheese spread's current purchasers are lost. That is, if we let p be the true proportion of all current purchasers who would stop buying the cheese spread if the new spout were used, profits will increase as long as p is less than .10.

Suppose that (after trying the new spout) 63 of 1,000 randomly selected purchasers say that they would stop buying the cheese spread if the new spout were used. To assess whether p is less than .10, we will assume for the sake of argument that p equals .10, and we will use the sample information to weigh the evidence against this assumption and in favor of the conclusion that p is less than .10. Let the random variable x represent the number of the 1,000 purchasers who say they would stop buying the cheese spread. Assuming that p equals .10, then x is a binomial random variable with $n = 1{,}000$ and $p = .10$. Because the sample result of 63 is less than $\mu = np = 1{,}000(.1) = 100$, the expected value of x when p equals .10, we have some evidence to contradict the assumption that p equals .10. To evaluate the strength of this evidence, we calculate the probability that *63 or fewer* of the 1,000 randomly selected purchasers would say that they would stop buying the cheese spread if the new spout were used if, in fact, p equals .10.

FIGURE 6.26 Approximating the Binomial Probability $P(x \leq 63)$ by Using the Normal Curve When $\mu = np = 100$ and $\sigma = \sqrt{npq} = 9.4868$



6.5 The Exponential Distribution (Optional)

- Suppose that some event occurs as a Poisson process
  - That is, the number of times an event occurs is a Poisson random variable
- Let x be the random variable of the interval between successive occurrences of the event
  - The interval can be some unit of time or space
- Then x is described by the exponential distribution
  - With parameter λ, which is the mean number of events that can occur per given interval

The Exponential Distribution Continued

$$f(x)=\begin{cases}\lambda e^{-\lambda x} & \text{for } x \ge 0 \\ 0 & \text{otherwise}\end{cases}$$

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

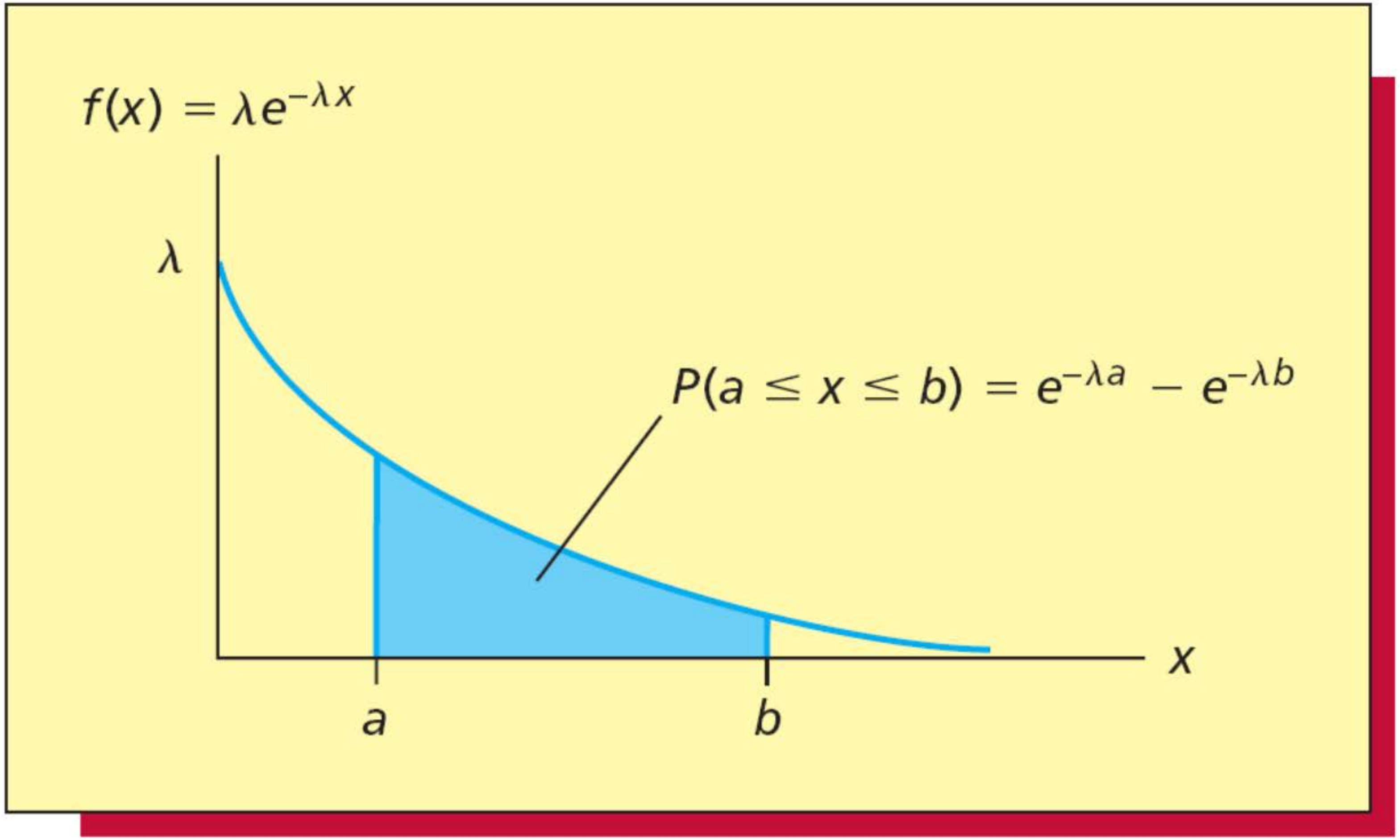
$$and$$

$$P(x \le c) = 1 - e^{-\lambda c} \text{ and } P(x \ge c) = e^{-\lambda c}$$

$$\mu_X = 1/\lambda \text{ and } \sigma_X = 1/\lambda$$

The Exponential Distribution Continued

Figure 6.27



Example

- λ = 20.8 errors per year
- λ = 0.4 errors per week
- Probability of one to two weeks

$$P(1 \le x \le 2) = e^{-\lambda a} - e^{-\lambda b} = e^{-\lambda(1)} - e^{-\lambda(2)}$$
$$= e^{-.4(1)} - e^{-.4(2)} = e^{-.4} - e^{-.8}$$
$$= .6703 - .4493 = .221$$

6.6 The Normal Probability Plot (Optional)

- A graphic used to visually check to see if sample data comes from a normal distribution
- A straight line indicates a normal distribution
- The more curved the line, the less normal the data is

Creating a Normal Probability Plot

1. Rank order the data from smallest to largest
2. For each data point, compute the value i/(n+1)
   - i is the data point's position on the list
3. For each data point, compute the standardized normal quantile value (O_i)
   - O_i is the z value that gives an area i/(n+1) to its left
4. Plot data points against O_i
5. Straight line indicates normal distribution

Sample Normal Probability Plots

Figures 6.30 to 6.32

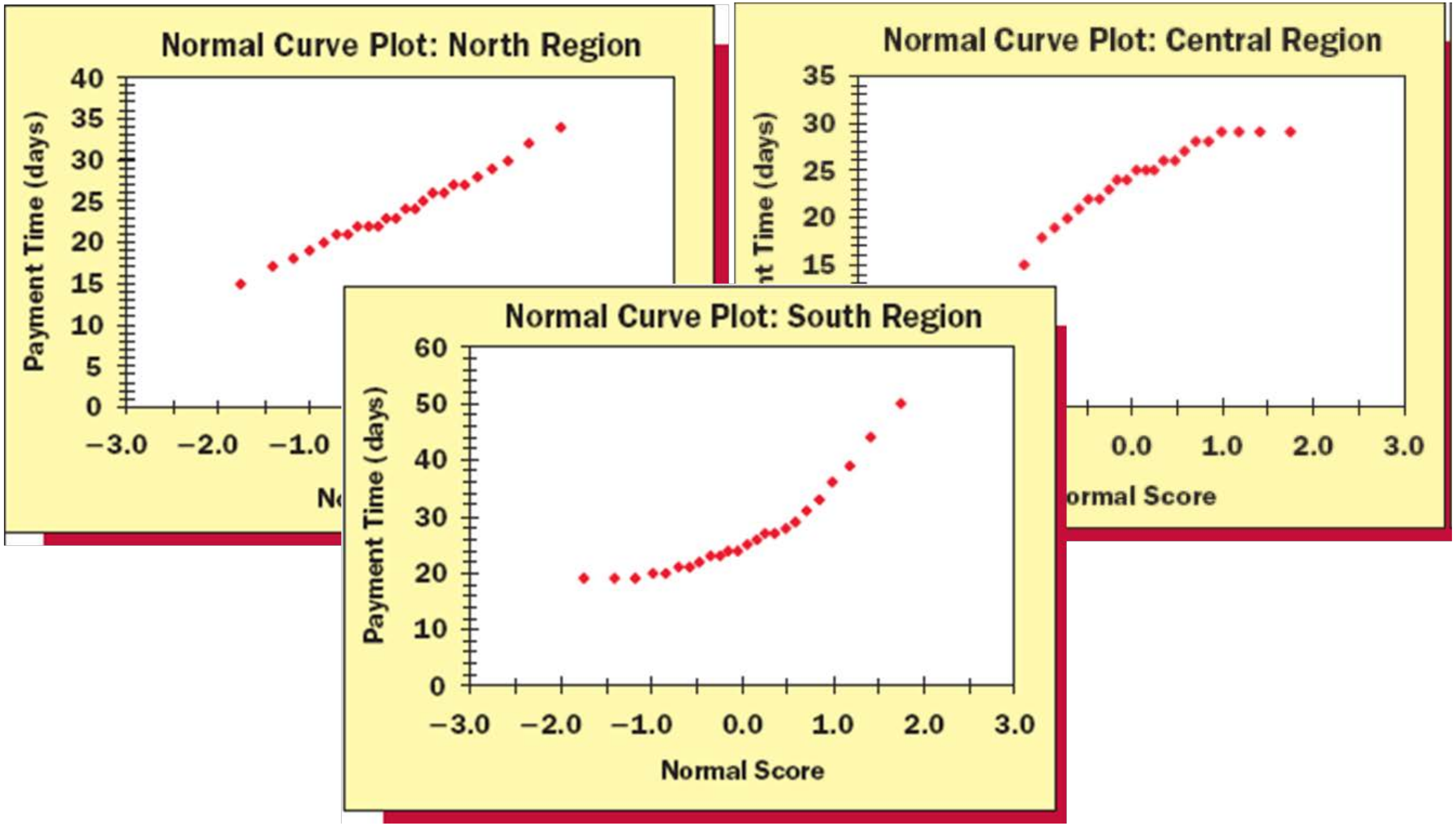


TABLE 6.3 Twenty-four Randomly Selected Payment Times for Each of Three Geographical Regions in the United States DS RegPayTime

North Region	Central Region	South Region
26	26	28
27	28	31
21	21	21
22	22	23
22	23	23
23	24	24
27	27	29
20	19	20
22	22	22
29	29	36
18	15	19
24	25	25
28	28	33
26	26	27
21	20	21
32	29	44
23	24	24
24	25	26
25	25	27
15	7	19
17	12	19
19	18	20
34	29	50
30	29	39

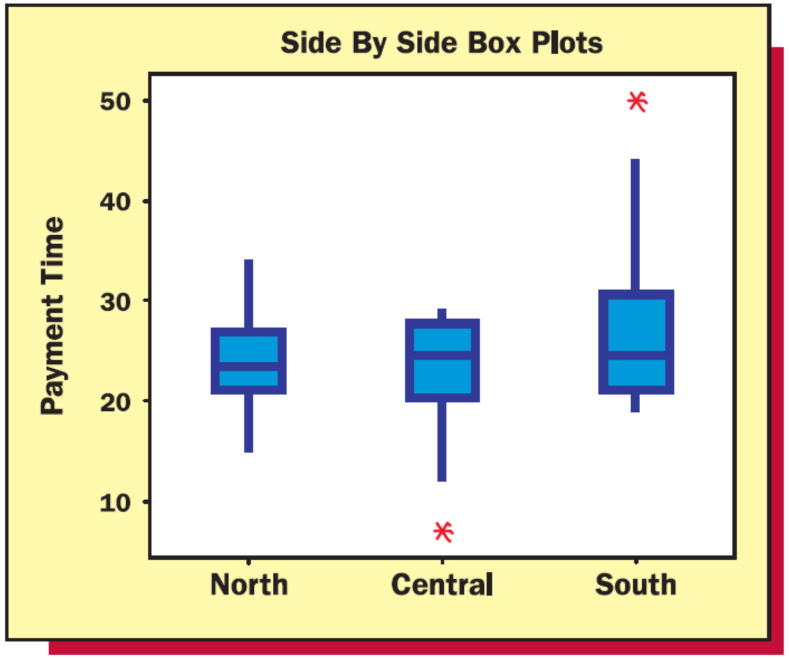


FIGURE 6.28 Dot Plot of the Payment Times for the North Region

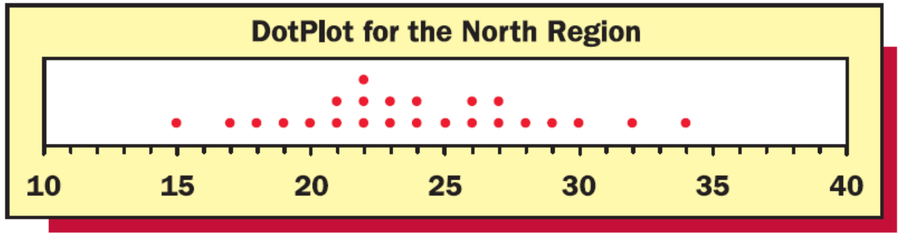
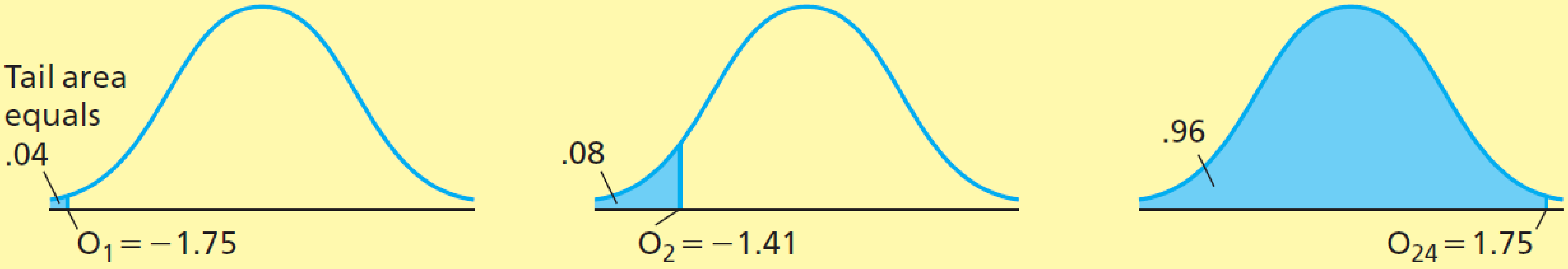


TABLE 6.4 Calculations for Normal Probability Plots in the e-billing Example

Ordered North Region Payment Times Column (1)	Observation Number (i) Column (2)	Area $i/(n + 1)$ Column (3)	z value O_i Column (4)	Ordered Central Region Payment Times Column (5)	Ordered South Region Payment Times Column (6)
15	1	0.04	−1.75	7	19
17	2	0.08	−1.41	12	19
18	3	0.12	−1.18	15	19
19	4	0.16	−0.99	18	20
20	5	0.2	−0.84	19	20
21	6	0.24	−0.71	20	21
21	7	0.28	−0.58	21	21
22	8	0.32	−0.47	22	22
22	9	0.36	−0.36	22	23
22	10	0.4	−0.25	23	23
23	11	0.44	−0.15	24	24
23	12	0.48	−0.05	24	24
24	13	0.52	0.05	25	25
24	14	0.56	0.15	25	26
25	15	0.6	0.25	25	27
26	16	0.64	0.36	26	27
26	17	0.68	0.47	26	28
27	18	0.72	0.58	27	29
27	19	0.76	0.71	28	31
28	20	0.8	0.84	28	33
29	21	0.84	0.99	29	36
30	22	0.88	1.18	29	39
32	23	0.92	1.41	29	44
34	24	0.96	1.75	29	50

FIGURE 6.29 Calculating Standardized Normal Quantile Values



(a) Calculating O_1 (b) Calculating O_2 (c) Calculating O_{24}

Summary

Chapter Summary

In this chapter we have discussed **continuous probability distributions.** We began by learning that **a continuous probability distribution is described by a continuous probability curve** and that in this context **probabilities are areas under the probability curve.** We next studied two important continuous probability distributions—**the uniform distribution** and **the normal distribution.** In particular, we concentrated on the normal distribution, which is the most important continuous probability distribution. We learned about the properties of the normal curve, and we saw how to use a **normal table** to find various areas under a normal curve. We then demonstrated how we can use a normal curve probability to make a statistical inference. We continued with an optional section that explained how we can use a normal curve probability to approximate a binomial probability. Then we presented an optional section that discussed another important continuous probability distribution—**the exponential distribution,** and we saw how this distribution is related to the Poisson distribution. Finally, we concluded this chapter with an optional section that explained how to use a **normal probability plot** to decide whether data come from a normal distribution.

Thank you!