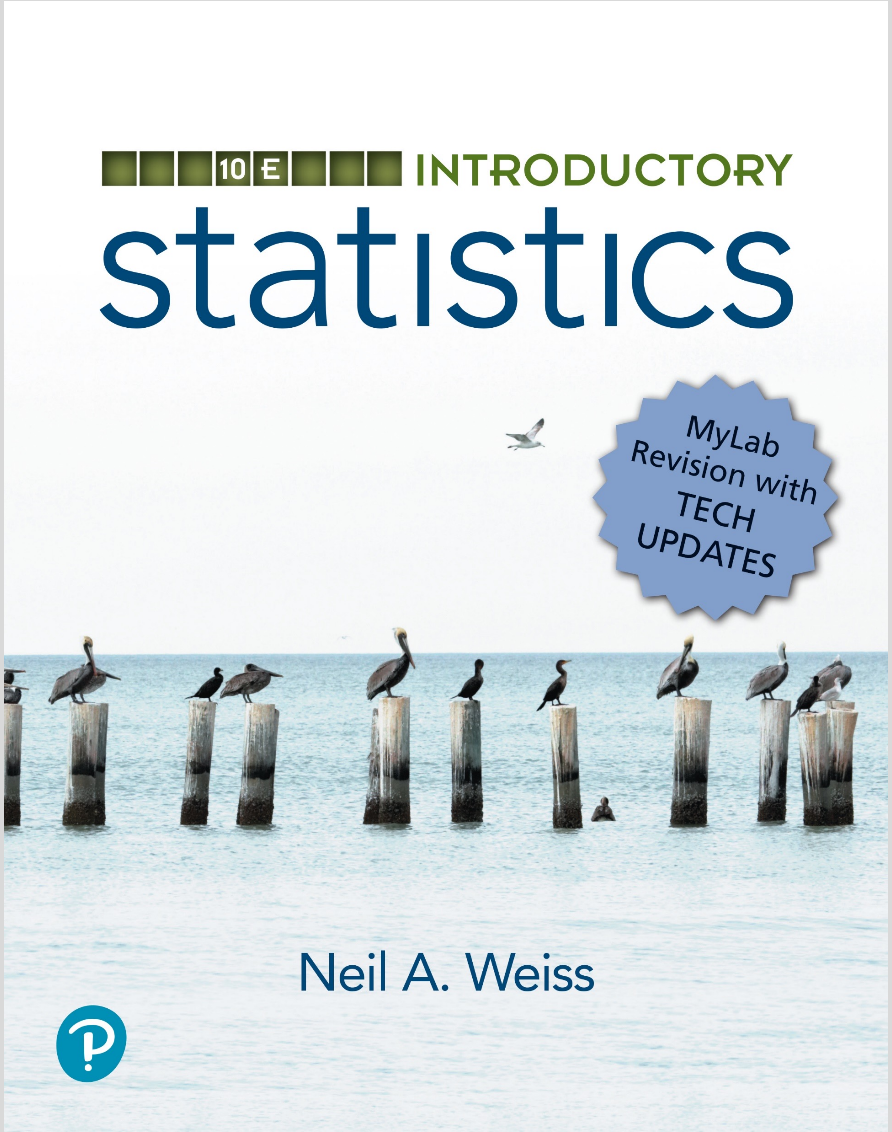
# Chapter 6

## The Normal Distribution



# Chapter 6

## The Normal Distribution

# Section 6.1

## Introducing Normally Distributed Variables

# Key Fact 6.1

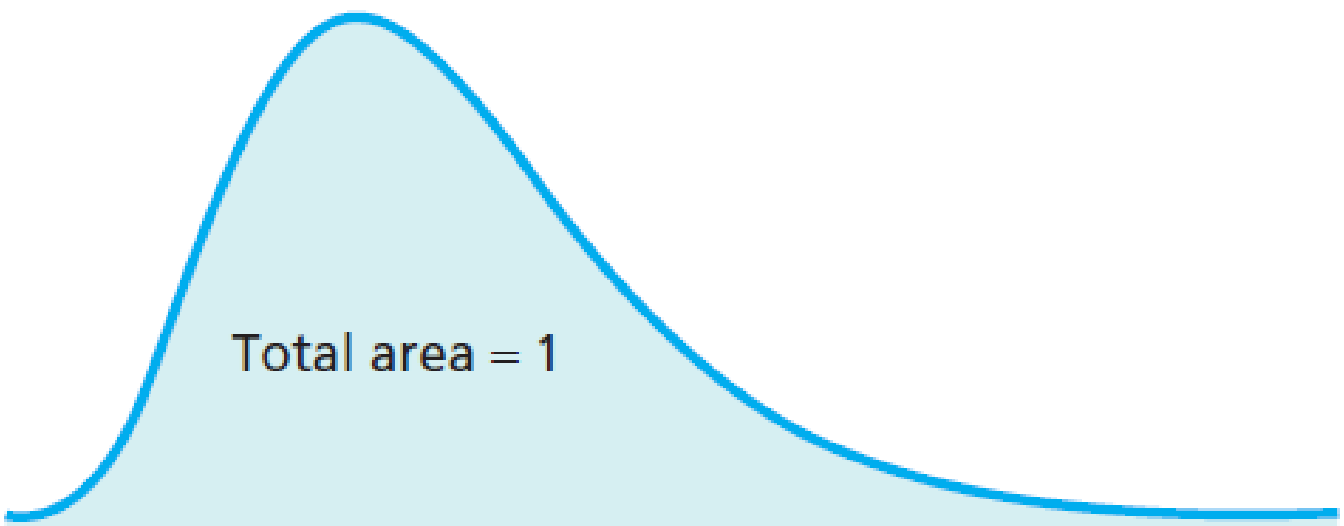
### Basic Properties of Density Curves

**Property 1:** A density curve is always on or above the horizontal axis.

**Property 2:** The total area under a density curve (and above the horizontal axis) equals 1.

# Figure 6.1

Properties 1 and 2 of Key Fact 6.1



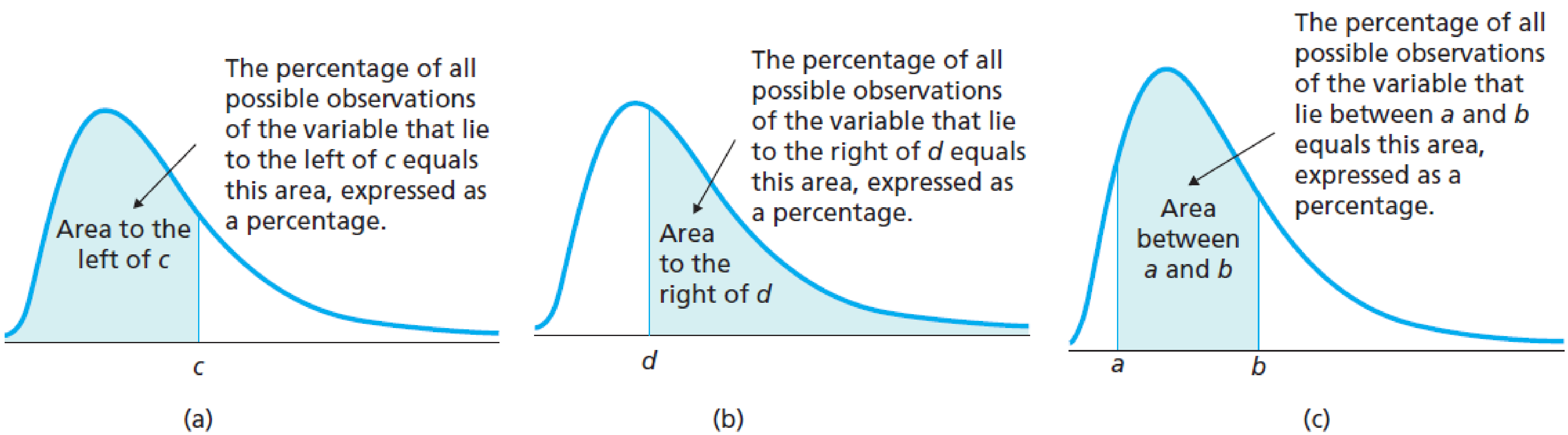
# Key Fact 6.2

**Variables and Their Density Curves**

For a variable with a density curve, the percentage of all possible observations of the variable that lie within any specified range equals (at least approximately) the corresponding area under the density curve, expressed as a percentage.

# Figure 6.2

## Illustration of Key Fact 6.2



# Definition 6.1

**Normally Distributed Variable**

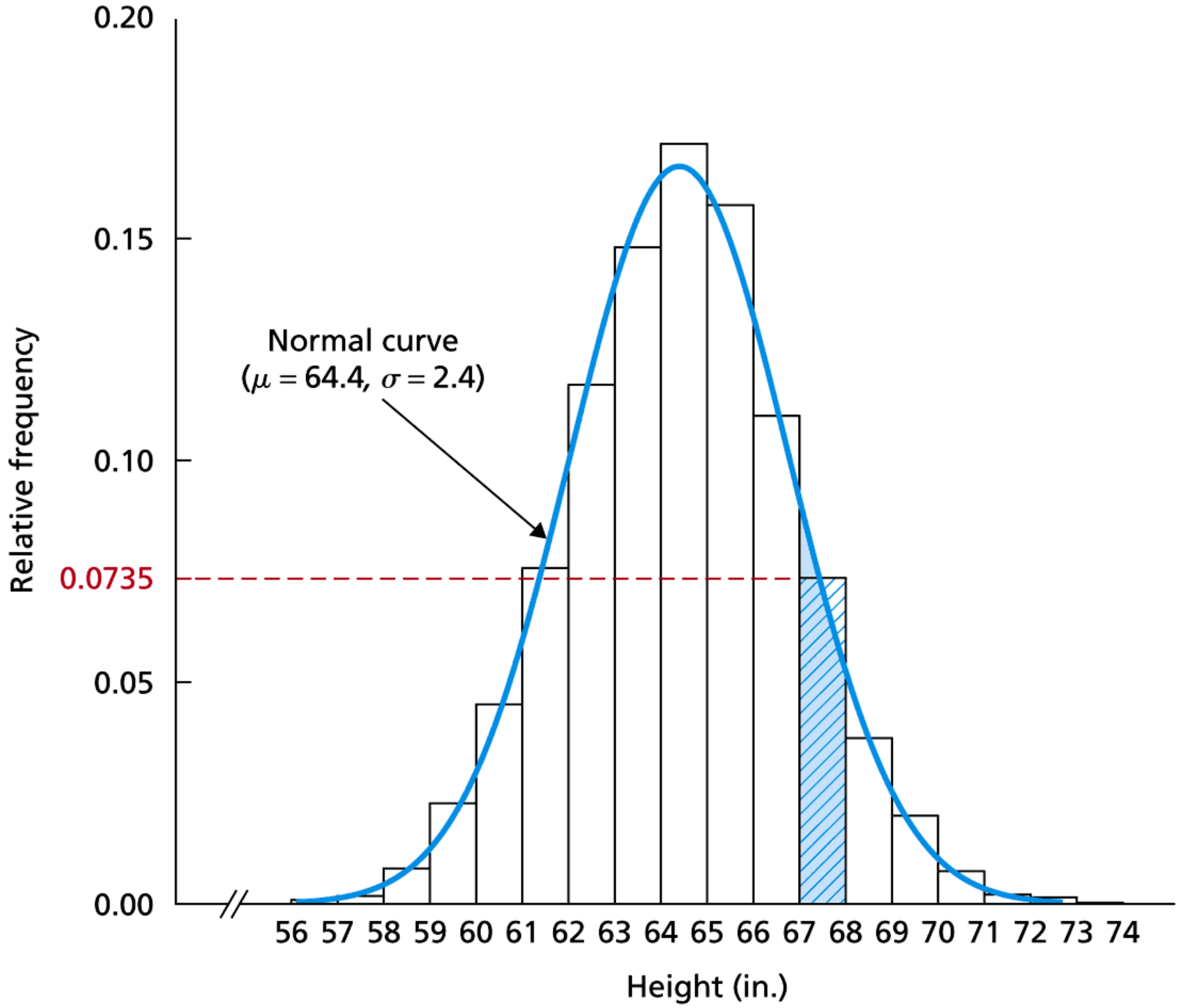
A variable is said to be a **normally distributed variable** or to have a **normal distribution** if its distribution has the shape of a normal curve.

# Table 6.1

Frequency and relative-frequency distributions for heights

| Height (in.) | Frequency $f$ | Relative frequency |
|---|---|---|
| 56–under 57 | 3 | 0.0009 |
| 57–under 58 | 6 | 0.0018 |
| 58–under 59 | 26 | 0.0080 |
| 59–under 60 | 74 | 0.0227 |
| 60–under 61 | 147 | 0.0450 |
| 61–under 62 | 247 | 0.0757 |
| 62–under 63 | 382 | 0.1170 |
| 63–under 64 | 483 | 0.1480 |
| 64–under 65 | 559 | 0.1713 |
| 65–under 66 | 514 | 0.1575 |
| 66–under 67 | 359 | 0.1100 |
| 67–under 68 | 240 | 0.0735 |
| 68–under 69 | 122 | 0.0374 |
| 69–under 70 | 65 | 0.0199 |
| 70–under 71 | 24 | 0.0074 |
| 71–under 72 | 7 | 0.0021 |
| 72–under 73 | 5 | 0.0015 |
| 73–under 74 | 1 | 0.0003 |
| | 3264 | 1.0000 |

# Figure 6.10 Relative-frequency histogram for heights with superimposed normal curve



# Key Fact 6.3

**Normally Distributed Variables and Normal-Curve Areas**

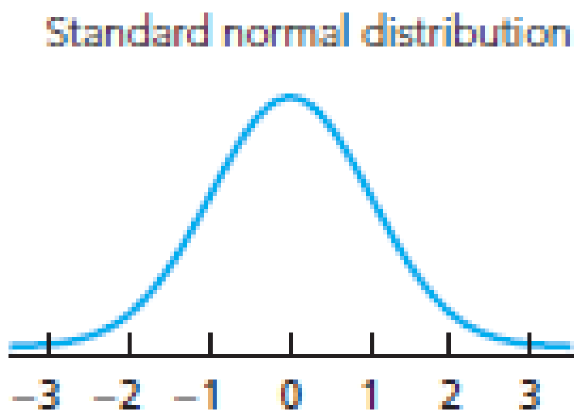
For a normally distributed variable, the percentage of all possible observations that lie within any specified range equals the corresponding area under its associated normal curve, expressed as a percentage. This result holds approximately for a variable that is approximately normally distributed.

# Definition 6.2

**Standard Normal Distribution; Standard Normal Curve**

A normally distributed variable having mean 0 and standard deviation 1 is said to have the **standard normal distribution.** Its associated normal curve is called the **standard normal curve**, which is shown in Fig. 6.11.

## Figure 6.11



# Key Fact 6.4

**Standardized Normally Distributed Variable**

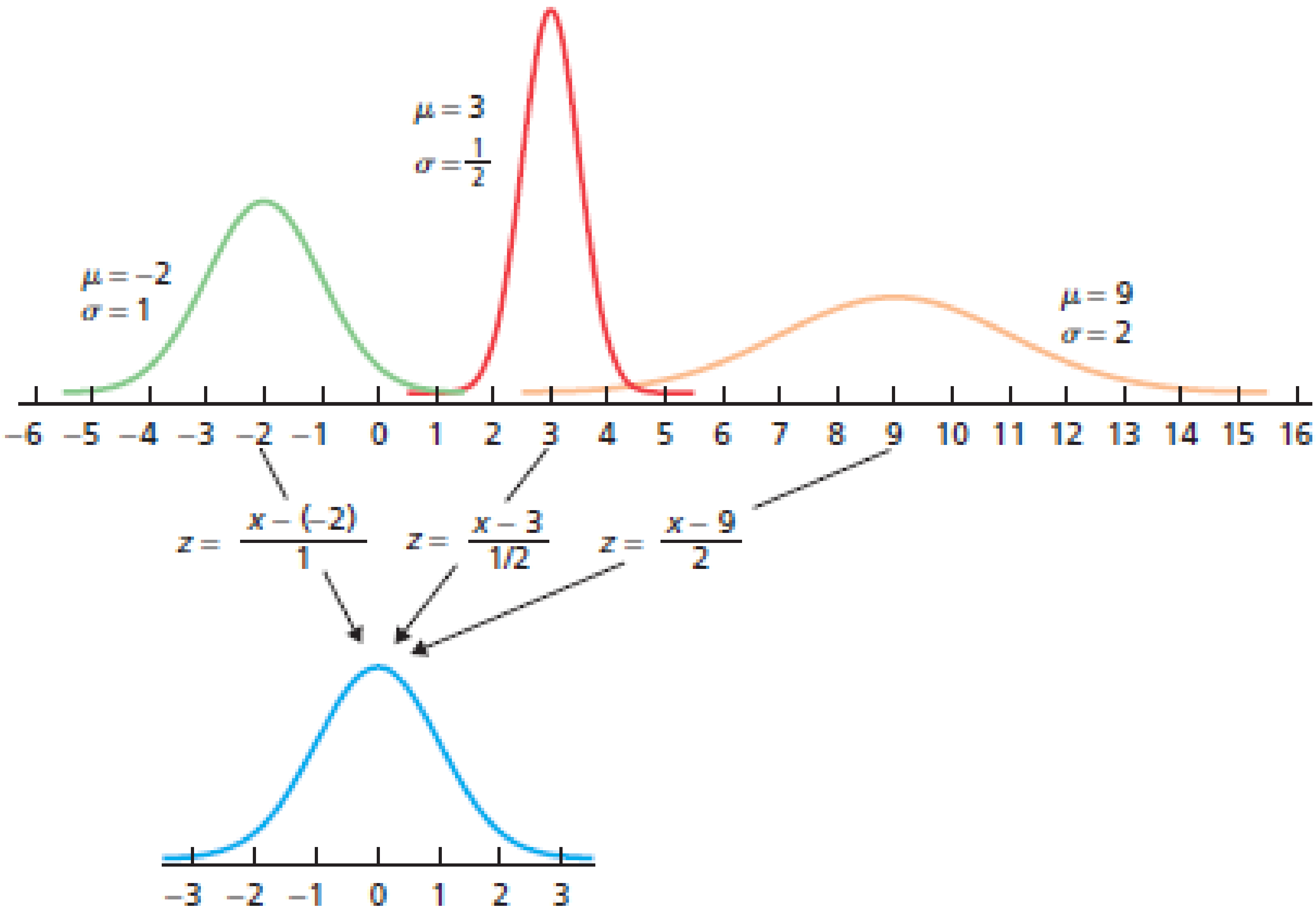
The standardized version of a normally distributed variable $x$,

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

has the standard normal distribution.

# Figure 6.12

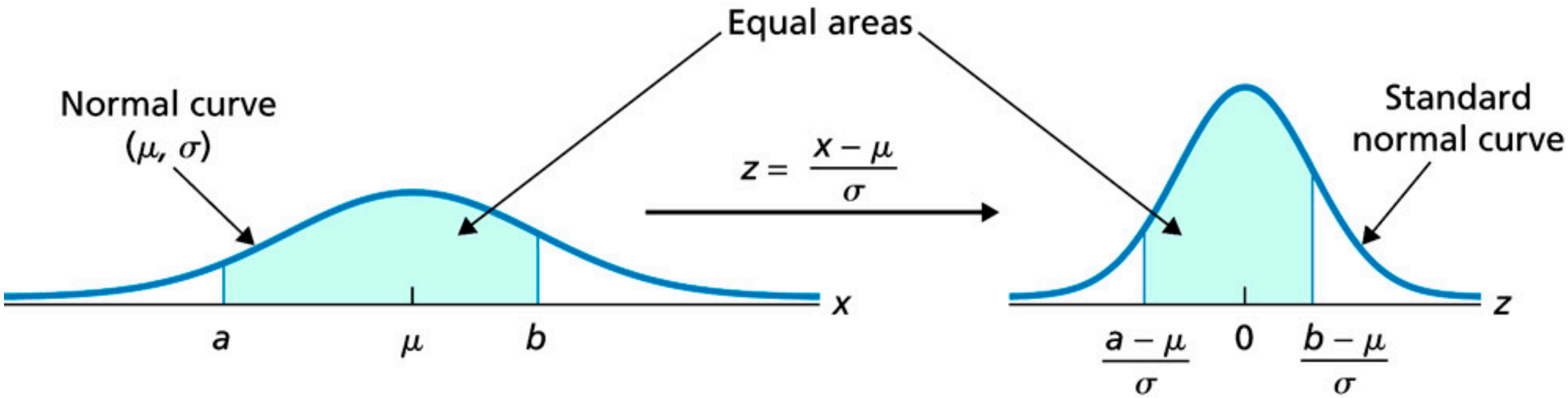
Standardizing normal distributions



# Figure 6.13

Finding percentages for a normally distributed variable from areas under the standard normal curve



# Section 6.2

## Areas Under the Standard Normal Curve

# Key Fact 6.5

## Basic Properties of the Standard Normal Curve

**Property 1:** The total area under the standard normal curve is 1.

**Property 2:** The standard normal curve extends indefinitely in both directions, approaching, but never touching, the horizontal axis as it does so.

**Property 3:** The standard normal curve is symmetric about 0; that is, the part of the curve to the left of the dashed line in Fig. 6.14 is the mirror image of the part of the curve to the right of it.

**Property 4:** Almost all the area under the standard normal curve lies between −3 and 3.

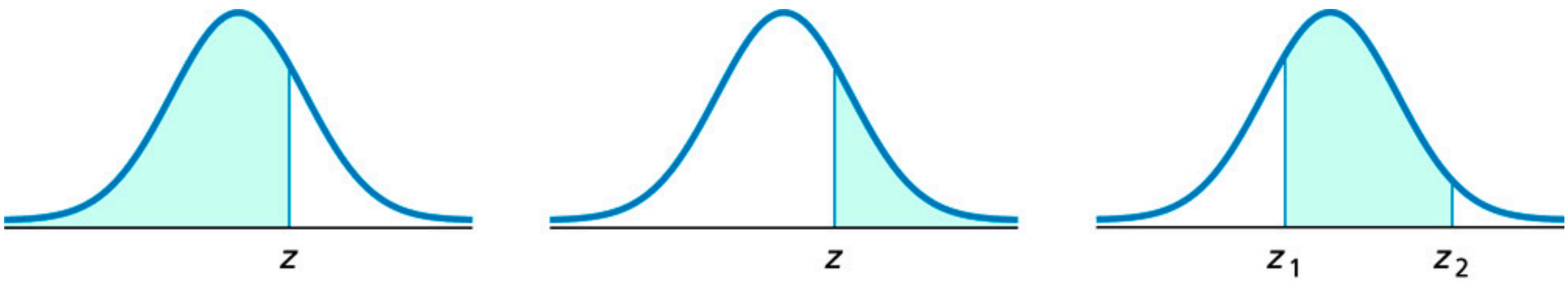
# Table 6.2

Areas under the standard normal curve

| | Second decimal place in *z* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *z* | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
| . | . | . | . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . | . | . | . |
| 1.0 | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| 1.1 | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8810 | 0.8830 |
| 1.2 | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| 1.3 | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| 1.4 | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| . | . | . | . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . | . | . | . |

# Figure 6.18

Using Table II to find the area under the standard normal curve that lies (a) to the left of a specified z-score, (b) to the right of a specified z-score, and (c) between two specified z-scores




(a) Shaded area: Area to left of $z$

(b) Shaded area: 1 – (Area to left of $z$)

(c) Shaded area: (Area to left of $z_2$) – (Area to left of $z_1$)

# Definition 6.3

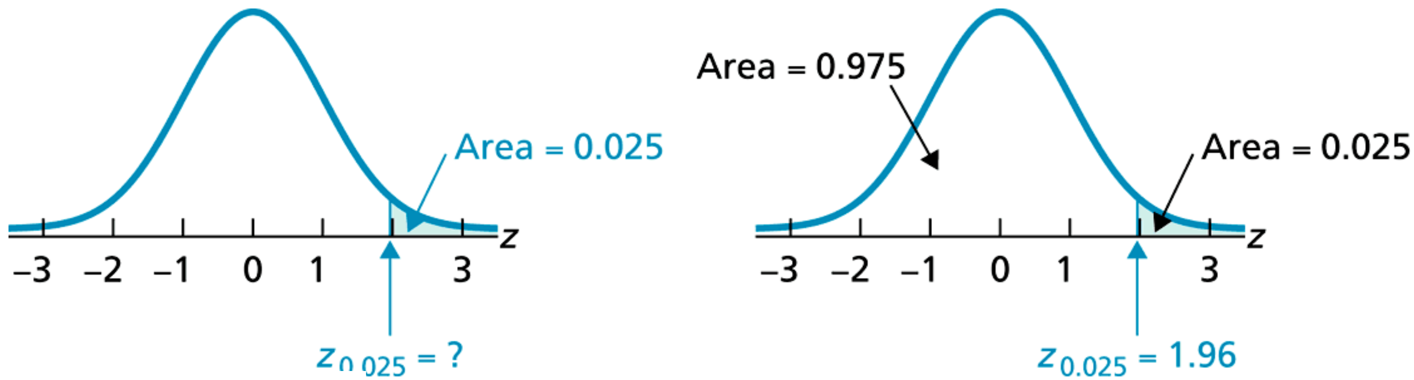
### The $z_\alpha$ Notation

The symbol $\mathbf{z}_\alpha$ is used to denote the $z$-score that has an area of $\alpha$ (alpha) to its right under the standard normal curve, as illustrated in Fig. 6.20. Read "$z_\alpha$" as "$z$ sub $\alpha$" or more simply as "$z\ \alpha$."

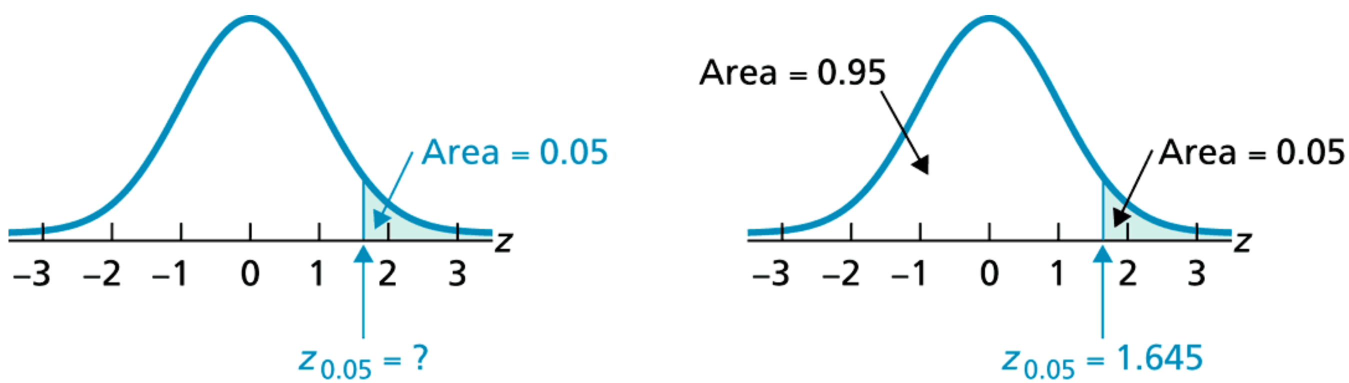
# Figures 6.21 & 6.22

Finding $z_{0.025}$




(b)

Finding $z_{0.05}$




(a) (b)

# Section 6.3

## Working with Normally Distributed Variables

# Procedure 6.1

## To Determine a Percentage or Probability for a Normally Distributed Variable

**Step 1** Sketch the normal curve associated with the variable.

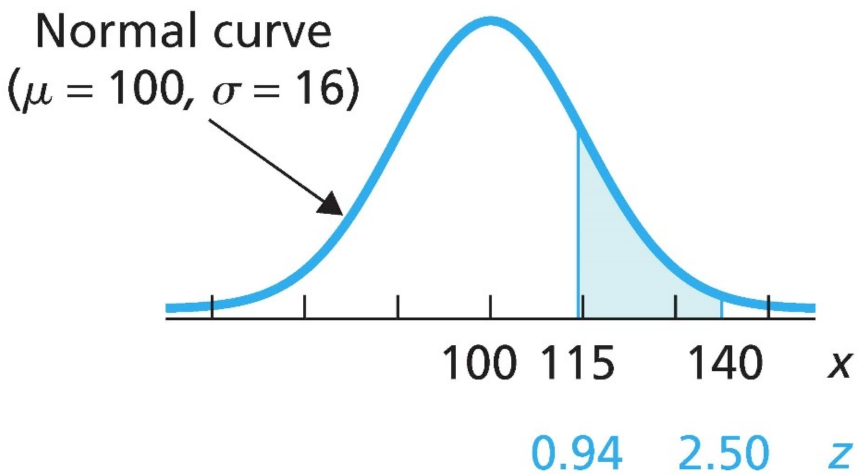
**Step 2** Shade the region of interest and mark its delimiting $x$-value(s).

**Step 3** Find the $z$-score(s) for the delimiting $x$-value(s) found in Step 2.

**Step 4** Use Table II to find the area under the standard normal curve delimited by the $z$-score(s) found in Step 3.

# Figure 6.25

Determination of the percentage of people having IQs between 115 and 140



# Key Fact 6.6

## Empirical Rule for Variables

For any variable whose distribution is bell-shaped (in particular, for any normally distributed variable), the following three properties hold.

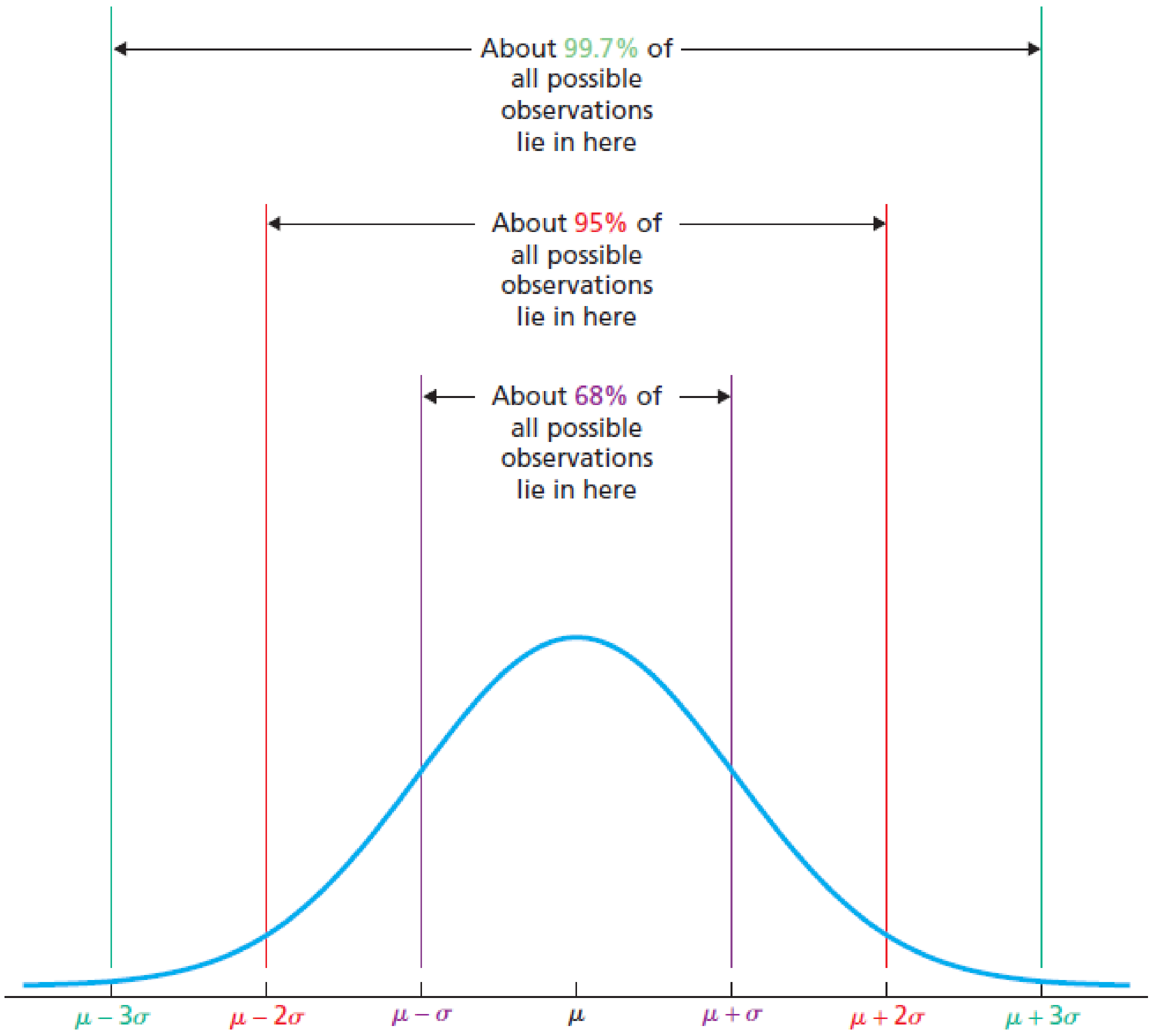
**Property 1:** Approximately 68% of all possible observations lie within one standard deviation to either side of the mean, that is, between $\mu - \sigma$ and $\mu + \sigma$.

**Property 2:** Approximately 95% of all possible observations lie within two standard deviations to either side of the mean, that is, between $\mu - 2\sigma$ and $\mu + 2\sigma$.

**Property 3:** Approximately 99.7% of all possible observations lie within three standard deviations to either side of the mean, that is, between $\mu - 3\sigma$ and $\mu + 3\sigma$.

These three properties are illustrated together in Fig. 6.26.

# Figure 6.26



About 99.7% of all possible observations lie in here
About 95% of all possible observations lie in here
About 68% of all possible observations lie in here
$\mu - 3\sigma$
$\mu - 2\sigma$
$\mu - \sigma$
$\mu$
$\mu + \sigma$
$\mu + 2\sigma$
$\mu + 3\sigma$

# Procedure 6.2

**To Determine the Observations Corresponding to a Specified Percentage or Probability for a Normally Distributed Variable**

**Step 1** Sketch the normal curve associated with the variable.

**Step 2** Shade the region of interest.

**Step 3** Use Table II to determine the $z$-score(s) delimiting the region found in Step 2.

**Step 4** Find the $x$-value(s) having the $z$-score(s) found in Step 3.

# Section 6.4

## Assessing Normality; Normal Probability Plots

# Key Fact 6.7

## Guidelines for Assessing Normality Using a Normal Probability Plot

To assess the normality of a variable using sample data, construct a normal probability plot.

- If the plot is roughly linear, you can assume that the variable is approximately normally distributed.
- If the plot is not roughly linear, you can assume that the variable is not approximately normally distributed.

These guidelines should be interpreted loosely for small samples but usually interpreted strictly for large samples.

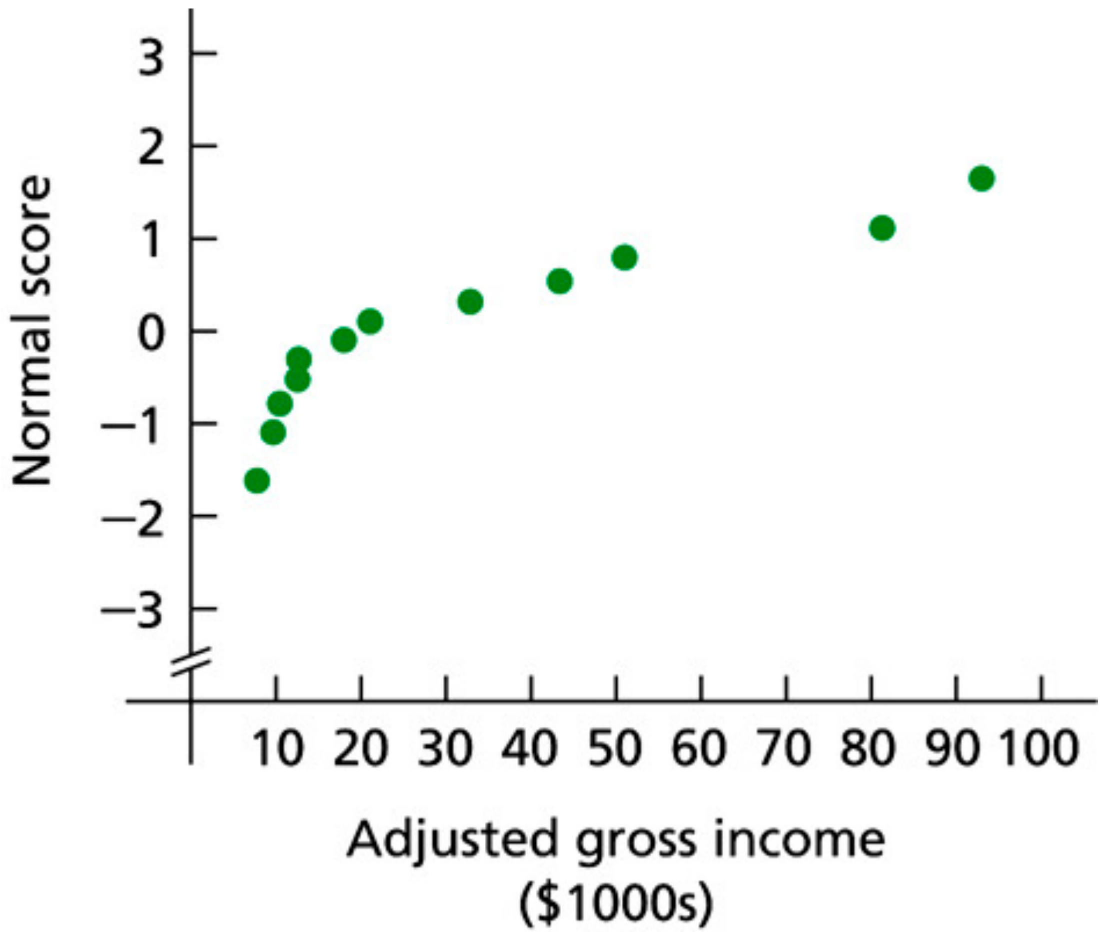
# Table 6.5

Ordered data and normal scores

| Adjusted gross income | Normal score |
|---|---|
| 7.8 | −1.64 |
| 9.7 | −1.11 |
| 10.6 | −0.79 |
| 12.7 | −0.53 |
| 12.8 | −0.31 |
| 18.1 | −0.10 |
| 21.2 | 0.10 |
| 33.0 | 0.31 |
| 43.5 | 0.53 |
| 51.1 | 0.79 |
| 81.4 | 1.11 |
| 93.1 | 1.64 |

# Figure 6.29

Normal probability plot for the sample of adjusted gross incomes

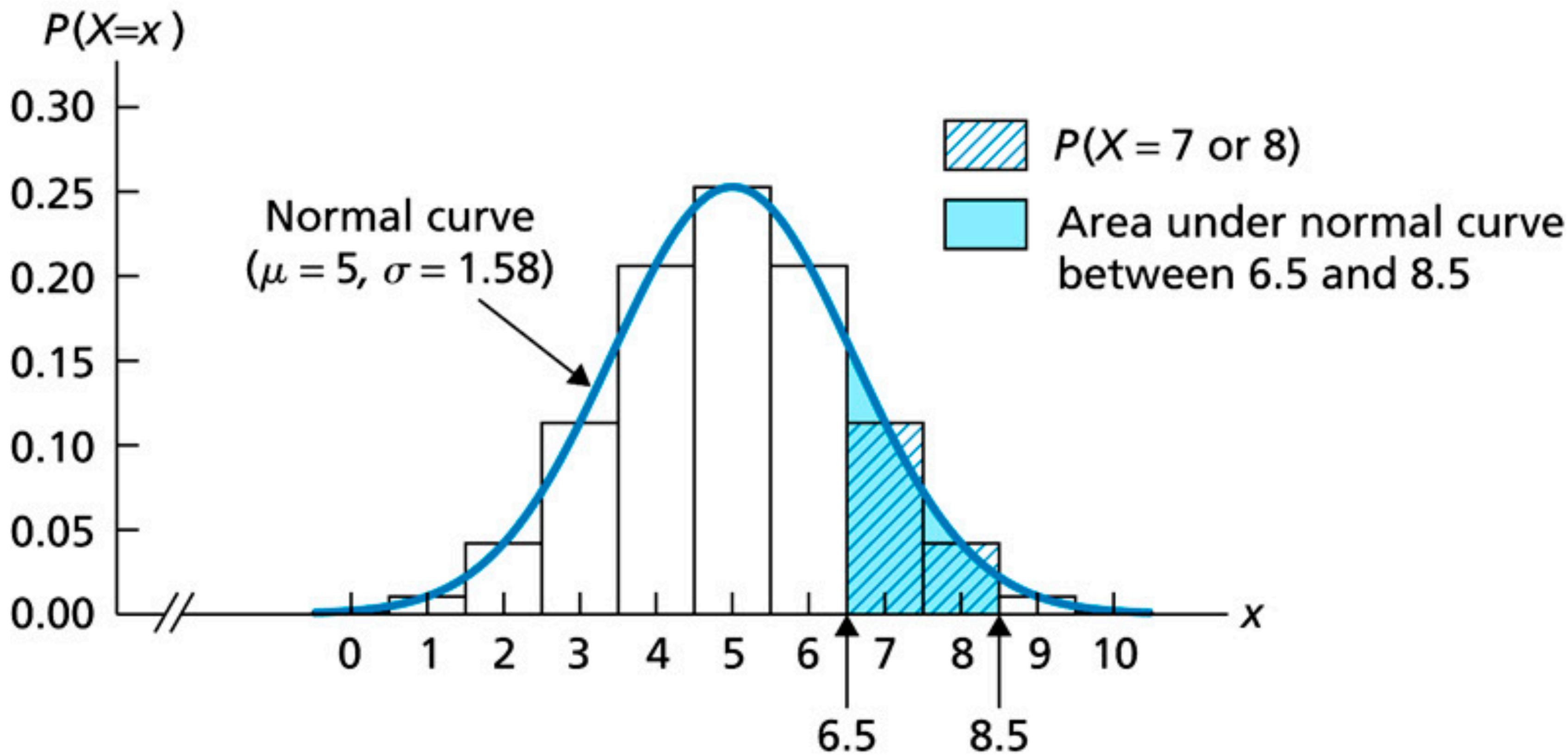


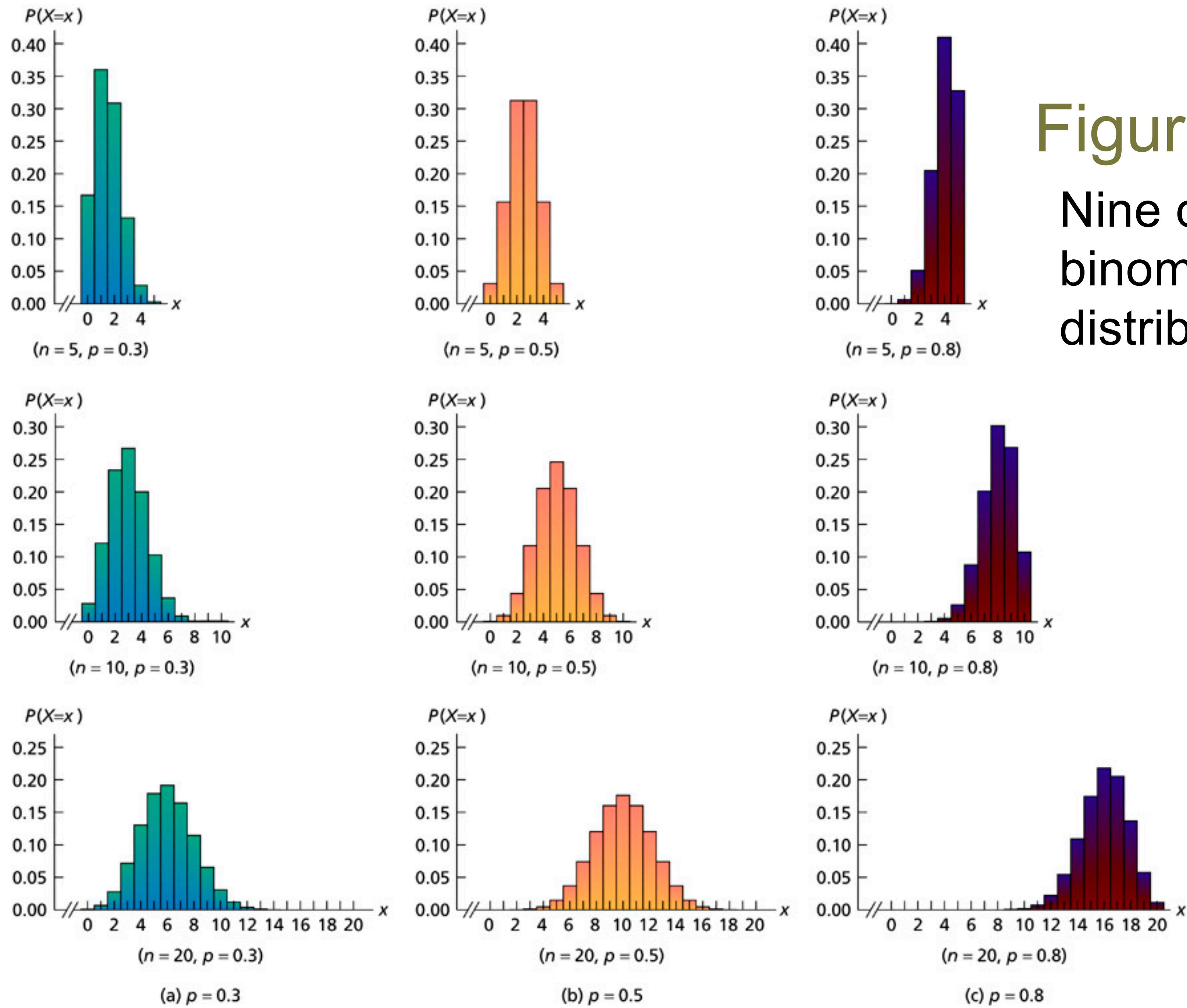
# Section 6.5

## Normal Approximation to the Binomial Distribution

# Figure 6.31

Probability histogram for *X* with superimposed normal curve






# Figure 6.33

Nine different binomial distributions

# Procedure 6.3

**To Approximate Binomial Probabilities by Normal-Curve Areas**

**Step 1** Find $n$, the number of trials, and $p$, the success probability.

**Step 2** Continue only if both $np$ and $n(1 - p)$ are 5 or greater.

**Step 3** Find $\mu$ and $\sigma$, using the formulas $\mu = np$ and $\sigma = \sqrt{np(1 - p)}$.

**Step 4** Make the correction for continuity, and find the required area under the normal curve with parameters $\mu$ and $\sigma$.