CHAPTER 5

Normal Probability Distributions

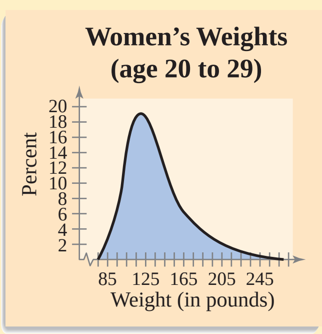
In 2000, the National Center for Health Statistics, located in Hyattsville, Maryland, began a 10-year program called *Healthy People 2010* to promote health through changes in people's lifestyles. It is too early to analyze the results of this program, but the results of a similar program that started in 1990, *Healthy People 2000*, are available. During the course of the program, some of the goals were met. For instance, heart disease and stroke death rates were down. Other goals were not met. For instance, although more adults were exercising, a quarter of all adults were still engaged in no physical activity.

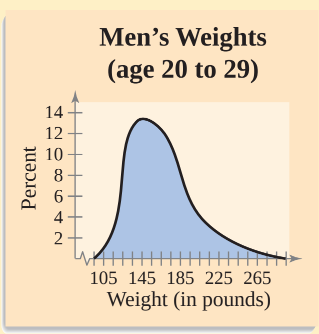
Where You've Been

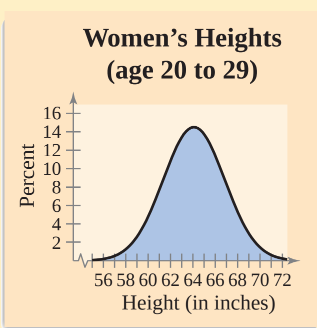
In Chapters 1 through 4, you learned how to collect and describe data, find the probability of an event, and analyze discrete probability distributions. You also learned that if a sample is used to make inferences about a population, then it is critical that the sample not be biased. Suppose, for instance, that you wanted to measure the serum cholesterol levels of adults in the United States. How would you organize the study? When the National Center for Health Statistics performed this study, it used random sampling and then classified the results according to the gender, ethnic background, and age of the participants. One conclusion from the study was that women's cholesterol levels tended to increase throughout their lives, whereas men's increased to age 65, and then decreased.

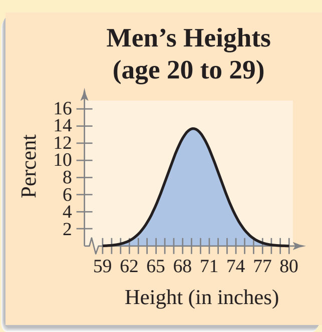
Where You're Going

In Chapter 5, you will learn how to recognize normal (bell-shaped) distributions and how to use their properties in real-life applications. Suppose that you worked for the U.S. National Center for Health Statistics and were collecting data about various physical traits of people in the United States. Which of the following would you expect to have bell-shaped, symmetric distributions: height, weight, cholesterol level, age, blood pressure, shoe size, reaction times, lung capacity? Of these, all except weight and age have distributions that are approximately normal. For instance, the four graphs below show the height and weight distributions for men and women in the United States aged 20 to 29. Notice that the height distributions are bell shaped, but the weight distributions are skewed right.









5.1 Introduction to Normal Distributions and the Standard Normal Distribution

What You Should Learn

- How to interpret graphs of normal probability distributions
- How to find areas under the standard normal curve

Properties of a Normal Distribution • The Standard Normal Distribution

Properties of a Normal Distribution

In Section 4.1, you learned that a **continuous random variable** has an infinite number of possible values that can be represented by an interval on the number line. Its probability distribution is called a **continuous probability distribution.** In this chapter, you will study the most important continuous probability distribution in statistics—the normal distribution. Normal distributions can be used to model many sets of measurements in nature, industry, and business. For instance, the systolic blood pressure of humans, the lifetime of television sets, and even housing costs are all normally distributed random variables.

Note to Instructor

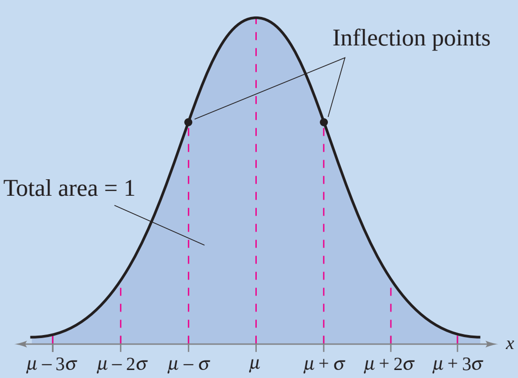
Draw several different continuous probability curves. Then point out that the normal (or Gaussian) curve is graphed using the formula shown at the bottom of the page. Have students discuss measures in nature that are normally distributed. Mention that often grades in a statistics class are not normally distributed.

GUIDELINES

Properties of a Normal Distribution

A **normal distribution** is a continuous probability distribution for a random variable x. The graph of a normal distribution is called the **normal curve.** A normal distribution has the following properties.

1. The mean, median, and mode are equal.
2. The normal curve is bell shaped and is symmetric about the mean.
3. The total area under the normal curve is equal to one.
4. The normal curve approaches, but never touches, the x-axis as it extends farther and farther away from the mean.
5. Between $\mu - \sigma$ and $\mu + \sigma$ (in the center of the curve) the graph curves downward. The graph curves upward to the left of $\mu - \sigma$ and to the right of $\mu + \sigma$. The points at which the curve changes from curving upward to curving downward are called *inflection points.*



Insight

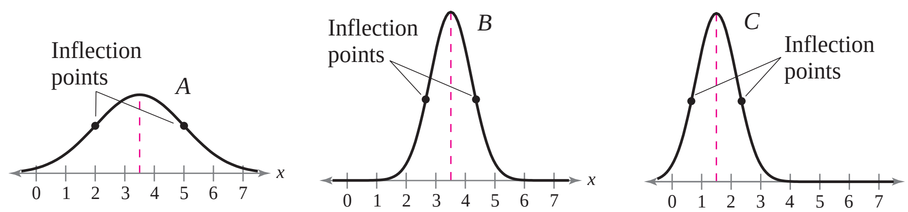
Because e and π are constants, a normal curve depends completely on two parameters, μ and σ.

If x is a continuous random variable having a normal distribution with mean μ and standard deviation σ, you can graph a normal curve using the equation

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}.$$

$e \approx 2.718$ and $\pi \approx 3.14$

A normal distribution can have any mean and any positive standard deviation. These two parameters, μ and σ, completely determine the shape of the normal curve. The mean gives the location of the line of symmetry, and the standard deviation describes how much the data are spread out.



Mean: $\mu = 3.5$ Standard deviation: $\sigma = 1.5$

Mean: $\mu = 3.5$ Standard deviation: $\sigma = 0.7$

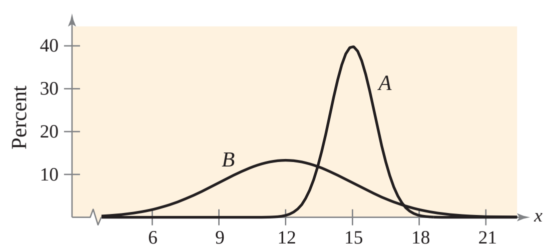
Mean: $\mu = 1.5$ Standard deviation: $\sigma = 0.7$

Notice that curve A and curve B above have the same mean, and curve B and curve C have the same standard deviation. The total area under each curve is 1.

EXAMPLE 1

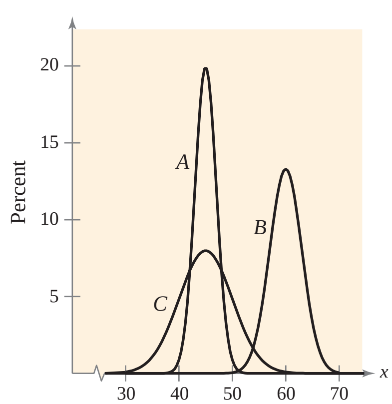
Understanding Mean and Standard Deviation

1. Which normal curve has a greater mean?
2. Which normal curve has a greater standard deviation?



SOLUTION

1. The line of symmetry of curve A occurs at $x = 15$. The line of symmetry of curve B occurs at $x = 12$. So, curve A has a greater mean.
2. Curve B is more spread out than curve A; so, curve B has a greater standard deviation.



Try It Yourself 1

Consider the normal curves shown at the left. Which normal curve has the greatest mean? Which normal curve has the greatest standard deviation? Justify your answers.

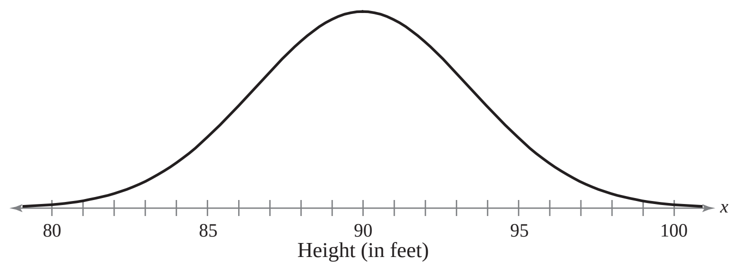
a. Find the location of the *line of symmetry* of each curve. Make a conclusion about which mean is greatest.
b. Determine which normal curve is *more spread out.* Make a conclusion about which standard deviation is greatest.

Answer: Page A35

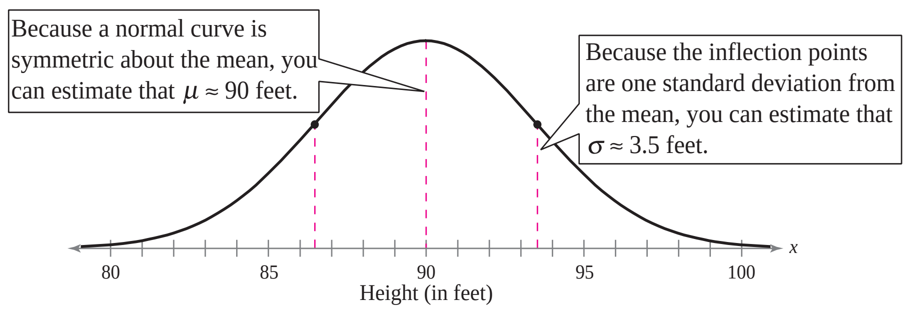
EXAMPLE 2

Interpreting Graphs of Normal Distributions

The heights (in feet) of fully grown white oak trees are normally distributed. The normal curve shown below represents this distribution. What is the mean height of a fully grown white oak tree? Estimate the standard deviation of this normal distribution.



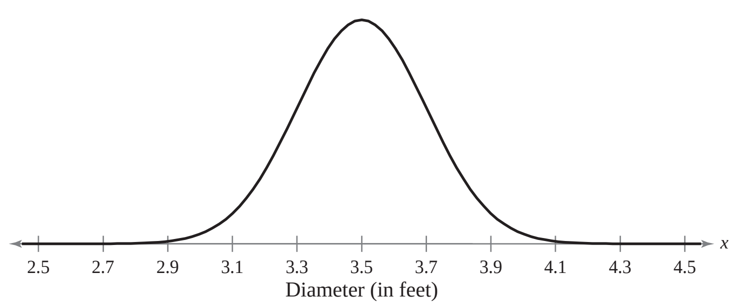
SOLUTION



Interpretation The heights of the oak trees are normally distributed with a mean of about 90 feet and a standard deviation of about 3.5 feet.

Try It Yourself 2

The diameters (in feet) of fully grown white oak trees are normally distributed. The normal curve shown below represents this distribution. What is the mean diameter of a fully grown white oak tree? Estimate the standard deviation of this normal distribution.



a. Find the *line of symmetry* and identify the mean.
b. Estimate the *inflection points* and identify the standard deviation.

Answer: Page A35

The Standard Normal Distribution

There are infinitely many normal distributions, each with its own mean and standard deviation. The normal distribution with a mean of 0 and a standard deviation of 1 is called **the standard normal distribution.** The horizontal scale of the graph of the standard normal distribution corresponds to z-scores. In Section 2.5, you learned that a z-score is a measure of position that indicates the number of standard deviations a value lies from the mean. Recall that you can transform an x-value to a z-score using the formula

$$z = \frac{\text{Value} - \text{Mean}}{\text{Standard deviation}}$$

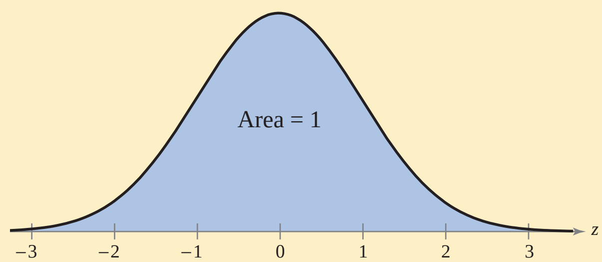
$$= \frac{x - \mu}{\sigma}. \qquad \text{Round to the nearest hundredth.}$$

Insight

Because every normal distribution can be transformed to the standard normal distribution, you can use z-scores and the standard normal curve to find areas (and therefore probability) under any normal curve.

DEFINITION

The **standard normal distribution** is a normal distribution with a mean of 0 and a standard deviation of 1.



Standard Normal Distribution

Note to Instructor

Mention that the formula for a normal probability density function on page 216 is greatly simplified when $\mu = 0$ and $\sigma = 1$.

$$y = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$

If each data value of a normally distributed random variable x is transformed into a z-score, the result will be the standard normal distribution. When this transformation takes place, the area that falls in the interval under the nonstandard normal curve is the *same* as that under the standard normal curve within the corresponding z-boundaries.

In Section 2.4, you learned to use the Empirical Rule to approximate areas under a normal curve when the values of the random variable x corresponded to −3, −2, −1, 0, 1, 2, or 3 standard deviations from the mean. Now, you will learn to calculate areas corresponding to other x-values. After you use the formula given above to transform an x-value to a z-score, you can use the Standard Normal Table in Appendix B. The table lists the cumulative area under the standard normal curve to the left of z for z-scores from −3.49 to 3.49. As you examine the table, notice the following.

Study Tip

It is important that you know the difference between x and z. The random variable x is sometimes called a raw score and represents values in a *nonstandard* normal distribution, whereas z represents values in the *standard* normal distribution.

Properties of the Standard Normal Distribution

1. The cumulative area is close to 0 for z-scores close to $z = -3.49$.
2. The cumulative area increases as the z-scores increase.
3. The cumulative area for $z = 0$ is 0.5000.
4. The cumulative area is close to 1 for z-scores close to $z = 3.49$.

Note to Instructor

If you prefer that your students use a 0-to-z table, refer them to Appendix A, where an alternative presentation for this material is given.

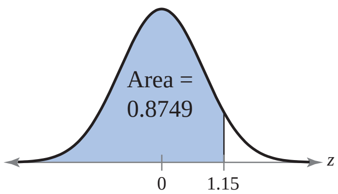
EXAMPLE 3

Using the Standard Normal Table

1. Find the cumulative area that corresponds to a z-score of 1.15.
2. Find the cumulative area that corresponds to a z-score of -0.24.

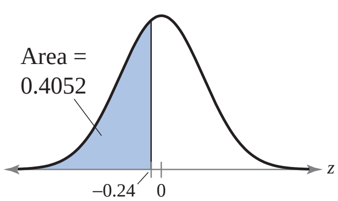
SOLUTION

1. Find the area that corresponds to $z = 1.15$ by finding 1.1 in the left column and then moving across the row to the column under 0.05. The number in that row and column is 0.8749. So, the area to the left of $z = 1.15$ is 0.8749.



z	.00	.01	.02	.03	.04	.05	.06
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279

2. Find the area that corresponds to $z = -0.24$ by finding -0.2 in the left column and then moving across the row to the column under 0.04. The number in that row and column is 0.4052. So, the area to the left of $z = -0.24$ is 0.4052.

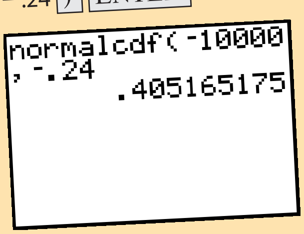


z	.09	.08	.07	.06	.05	.04	.03
−3.4	.0002	.0003	.0003	.0003	.0003	.0003	.0003
−3.3	.0003	.0004	.0004	.0004	.0004	.0004	.0004
−3.2	.0005	.0005	.0005	.0006	.0006	.0006	.0006
−0.5	.2776	.2810	.2843	.2877	.2912	.2946	.2981
−0.4	.3121	.3156	.3192	.3228	.3264	.3300	.3336
−0.3	.3483	.3520	.3557	.3594	.3632	.3669	.3707
−0.2	.3859	.3897	.3936	.3974	.4013	.4052	.4090
−0.1	.4247	.4286	.4325	.4364	.4404	.4443	.4483
−0.0	.4641	.4681	.4721	.4761	.4801	.4840	.4880

Study Tip

You can use a computer or calculator to find the cumulative area that corresponds to a z-score. For instance, here are instructions for finding the area that corresponds to $z = -0.24$ on a TI-83.

2nd DISTR 2 −10000, −.24) ENTER



Try It Yourself 3

1. Find the area under the curve to the left of a z-score of -2.19.
2. Find the area under the curve to the left of a z-score of 2.17.

a. Locate the given z-score and *find the area* that corresponds to it in the Standard Normal Table. *Answer: Page A36*

When the z-score is not in the table, use the entry closest to it. If the given z-score is exactly midway between two z-scores, then use the area midway between the corresponding areas.

You can use the following guidelines to find various types of areas under the standard normal curve.

Note to Instructor

Students find these three options easy to work with. If you have previously used a 0-to-z table, you will appreciate that students never need be confused as to whether to add 0.5, subtract it from 0.5, or use the table entry to find a required probability.

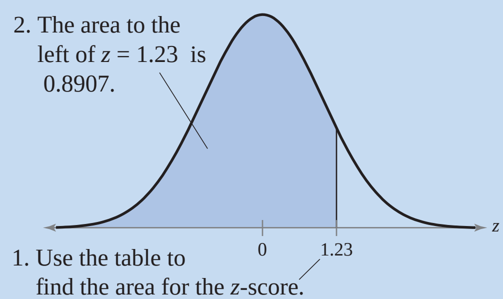
GUIDELINES

Finding Areas Under the Standard Normal Curve

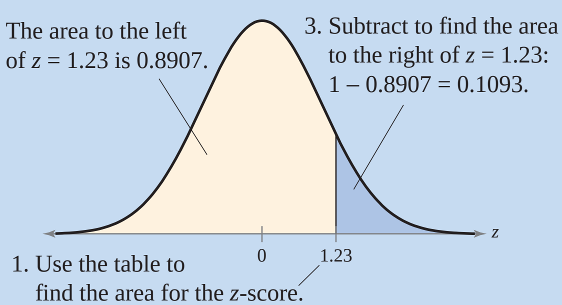
1. Sketch the standard normal curve and shade the appropriate area under the curve.

2. Find the area by following the directions for each case shown.

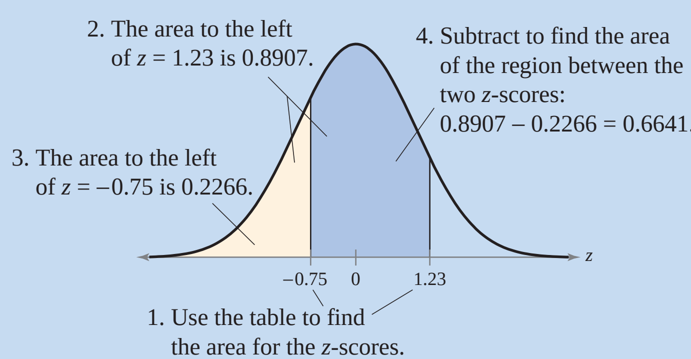
a. To find the area to the *left* of z, find the area that corresponds to z in the Standard Normal Table.



b. To find the area to the *right* of z, use the Standard Normal Table to find the area that corresponds to z. Then subtract the area from 1.



c. To find the area *between* two z-scores, find the area corresponding to each z-score in the Standard Normal Table. Then subtract the smaller area from the larger area.

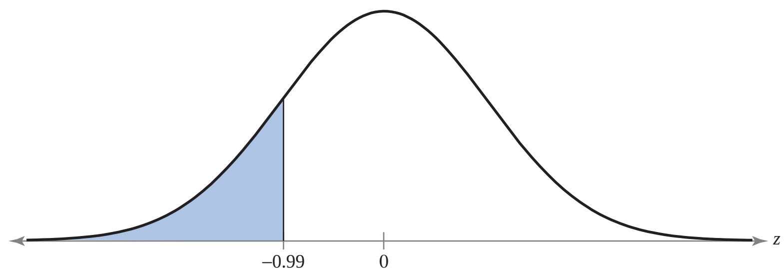


EXAMPLE 4

Finding Area Under the Standard Normal Curve

Find the area under the standard normal curve to the left of $z = -0.99$.

SOLUTION The area under the standard normal curve to the left of $z = -0.99$ is shown.



From the Standard Normal Table, this area is equal to 0.1611.

Try It Yourself 4

Find the area under the standard normal curve to the left of $z = 2.13$.

a. *Draw* the standard normal curve and shade the area under the curve and to the left of $z = 2.13$.

b. Use the Standard Normal Table to *find the area* that corresponds to $z = 2.13$.

Answer: Page A36

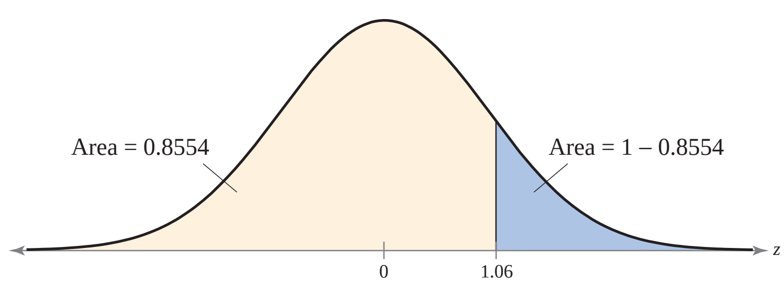
> **Insight**
>
> Because the normal distribution is a continuous probability distribution, the area under the standard normal curve to the left of a z-score gives the probability that z is less than that z-score. For instance, in Example 4, the area to the left of $z = -0.99$ is 0.1611. So, $P(z < -0.99) = 0.1611$, which is read as "the probability that z is less than -0.99 is 0.1611."

EXAMPLE 5

Finding Area Under the Standard Normal Curve

Find the area under the standard normal curve to the right of $z = 1.06$.

SOLUTION The area under the standard normal curve to the right of $z = 1.06$ is shown.



From the Standard Normal Table, the area to the left of $z = 1.06$ is 0.8554. Because the total area under the curve is 1, the area to the right of $z = 1.06$ is

$$\begin{aligned}\text{Area} &= 1 - 0.8554\\ &= 0.1446.\end{aligned}$$

Try It Yourself 5

Find the area under the standard normal curve to the right of $z = -2.16$.

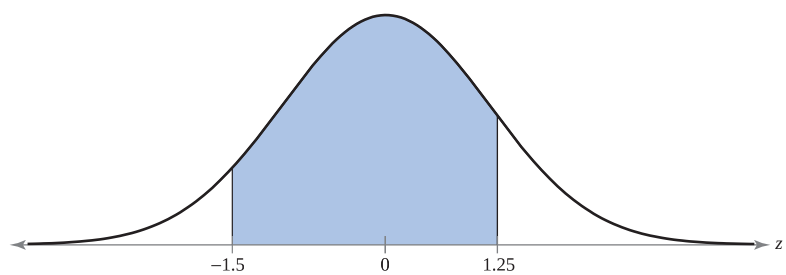
a. *Draw* the standard normal curve and shade the area below the curve and to the right of $z = -2.16$.
b. Use the Standard Normal Table to *find the area* to the left of $z = -2.16$.
c. *Subtract* the area from 1. *Answer: Page A36*

EXAMPLE 6

Finding Area Under the Standard Normal Curve

Find the area under the standard normal curve between $z = -1.5$ and $z = 1.25$.

SOLUTION The area under the standard normal curve between $z = -1.5$ and $z = 1.25$ is shown.



From the Standard Normal Table, the area to the left of $z = 1.25$ is 0.8944 and the area to the left of $z = -1.5$ is 0.0668. So, the area between $z = -1.5$ and $z = 1.25$ is

$$\text{Area} = 0.8944 - 0.0668$$
$$= 0.8276.$$

Interpretation So, 82.76% of the area under the curve falls between $z = -1.5$ and $z = 1.25$.

Try It Yourself 6

Find the area under the standard normal curve between $z = -2.16$ and $z = -1.35$.

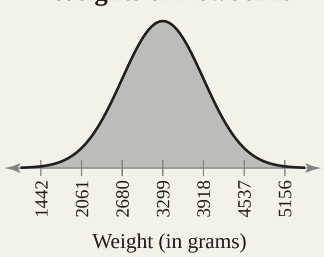
a. Use the Standard Normal Table to *find the area* to the left of $z = -1.35$.
b. Use the Standard Normal Table to *find the area* to the left of $z = -2.16$.
c. Subtract the smaller area from the larger area. *Answer: Page A36*

Recall in Section 2.5 you learned, using the Empirical Rule, that values lying more than two standard deviations from the mean are considered unusual. Values lying more than three standard deviations from the mean are considered *very* unusual. So if a z-score is greater than 2 or less than -2, it is unusual. If a z-score is greater than 3 or less than -3, it is *very* unusual.

Picturing the World

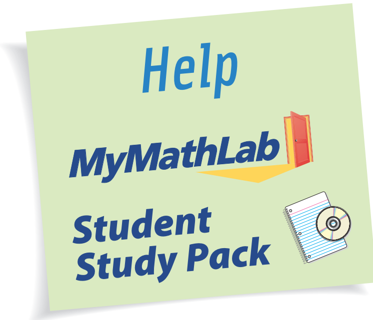
Each year the Centers for Disease Control and Prevention and the National Center for Health Statistics jointly publish a report summarizing the vital statistics from the previous year. According to one publication, the number of births in a recent year was 4,021,726. The weights of the newborns can be approximated by a normal distribution, as shown by the following graph.

Weights of Newborns



The weights of three newborns are 2000 grams, 3000 grams, and 4000 grams. Find the z-score that corresponds to each weight. Are any of these unusually heavy or light?

5.1 Exercises

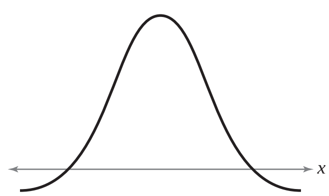


Building Basic Skills and Vocabulary

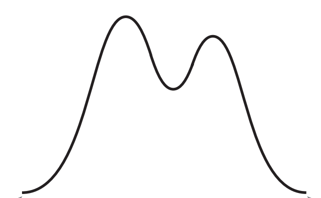
1. Find three real-life examples of a continuous variable. Which do you think may be normally distributed? Why?
2. What is the total area under the normal curve?
3. Draw two normal curves that have the same mean but different standard deviations. Describe the similarities and differences.
4. Draw two normal curves that have different means but the same standard deviations. Describe the similarities and differences.
5. What is the mean of the standard normal distribution? What is the standard deviation of the standard normal distribution?
6. Describe how you can transform a nonstandard normal distribution to a standard normal distribution.
7. ***Getting at the Concept*** Why is it correct to say "a" normal distribution and "the" standard normal distribution?
8. ***Getting at the Concept*** If a z-score is zero, which of the following must be true? Explain your reasoning.
   (a) The mean is zero.
   (b) The corresponding x-value is zero.
   (c) The corresponding x-value is equal to the mean.

Graphical Analysis In Exercises 9–14, determine whether the graph could represent a variable with a normal distribution. Explain your reasoning.

9.

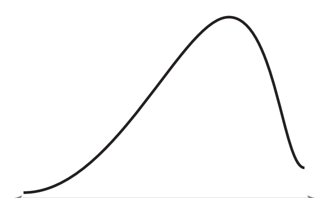


10.

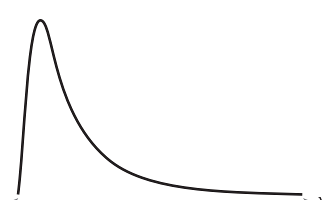


11.

12.



13.



14.

1. Answers will vary.
2. 1
3. Answers will vary.
   Similarities: The two curves will have the same line of symmetry.
   Differences: One curve will be more spread out than the other.
4. Answers will vary.
   Similarities: The two curves will have the same shape (i.e., equal standard deviations).
   Differences: The two curves will have different lines of symmetry.
5. $\mu = 0, \sigma = 1$
6. Transform each data value x into a z-score. This is done by subtracting the mean from x and dividing by the standard deviation. In symbols, $z = \dfrac{x - \mu}{\sigma}$.
7. "The" standard normal distribution is used to describe one specific normal distribution ($\mu = 0, \sigma = 1$). "A" normal distribution is used to describe a normal distribution with any mean and standard deviation.
8. (c) is true because a z-score equal to zero indicates that the corresponding x-value is equal to the mean.
9. No, the graph crosses the x-axis.
10. No, the graph is not symmetric.
11. Yes, the graph fulfills the properties of the normal distribution.
12. No, the graph is skewed left.
13. No, the graph is skewed right.
14. No, the graph is not bell shaped.

15. It is normal because it is bell shaped and symmetric.

16. It is skewed to the right. So it is not a normal distribution.

17. 0.3849

18. 0.4878

19. 0.6247

20. 0.0228

21. 0.9382

22. 0.5987

23. 0.975

24. 0.8997

25. 0.8289

26. 0.9599

27. 0.1003

28. 0.0099

29. 0.005

30. 0.0010

31. 0.05

32. 0.006

33. 0.475

34. 0.499

35. 0.437

36. 0.195

37. 0.95

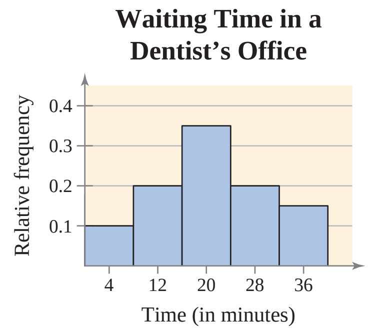
38. 0.9802

39. 0.2006

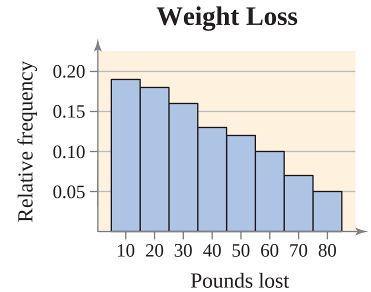
40. 0.05

Graphical Analysis In Exercises 15 and 16, determine whether the histogram represents data with a normal distribution. Explain your reasoning.

15.

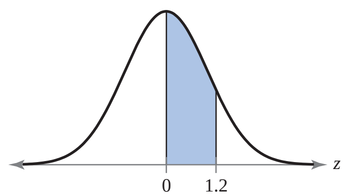


16.

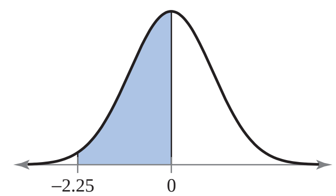


Graphical Analysis In Exercises 17–20, find the area of the indicated region under the standard normal curve.

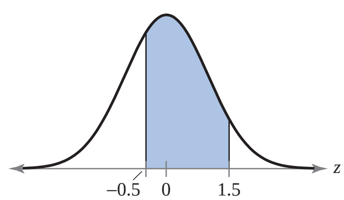
17.



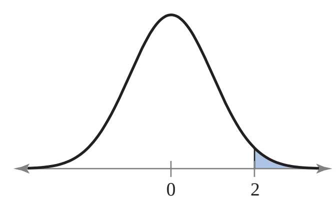
18.



19.



20.



Finding Area In Exercises 21–40, find the indicated area under the standard normal curve.

21. To the left of $z = 1.54$

22. To the left of $z = 0.25$

23. To the left of $z = 1.96$

24. To the left of $z = 1.28$

25. To the right of $z = -0.95$

26. To the right of $z = -1.75$

27. To the right of $z = 1.28$

28. To the right of $z = 2.33$

29. To the left of $z = -2.575$

30. To the left of $z = -3.08$

31. To the right of $z = 1.645$

32. To the right of $z = 2.51$

33. Between $z = 0$ and $z = 1.96$

34. Between $z = 0$ and $z = 3.09$

35. Between $z = -1.53$ and $z = 0$

36. Between $z = -0.51$ and $z = 0$

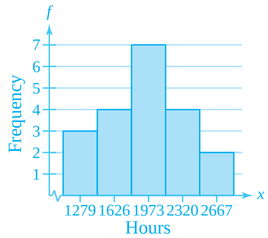
37. Between $z = -1.96$ and $z = 1.96$

38. Between $z = -2.33$ and $z = 2.33$

39. To the left of $z = -1.28$ or to the right of $z = 1.28$

40. To the left of $z = -1.96$ or to the right of $z = 1.96$

41. (a)

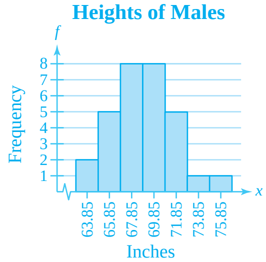


It is reasonable to assume that the life span is normally distributed because the histogram is nearly symmetric and bell shaped.

(b) 1941.35, 432.385

(c) The sample mean of 1941.35 hours is less than the claimed mean, so, on average, the bulbs in the sample lasted for a shorter time. The sample standard deviation of 432 hours is greater than the claimed standard deviation, so the bulbs in the sample had a greater variation in life span than the manufacturer's claim.

42. (a)



It is reasonable to assume that the heights are normally distributed because the histogram is nearly symmetric and bell shaped.

(b) 68.75, 2.847

(c) The mean of your sample is 0.45 inch less than that of the previous study, so the average height from the sample is less than in the previous study. The standard deviation is about 0.05 inch less than that of the previous study, so the heights are slightly less spread out than in the previous study.

Using and Interpreting Concepts

DATA **41.** ***Manufacturer Claims*** You work for a consumer watchdog publication and are testing the advertising claims of a light bulb manufacturer. The manufacturer claims that the life span of the bulb is normally distributed, with a mean of 2000 hours and a standard deviation of 250 hours. You test 20 light bulbs and get the following life spans.

2210, 2406, 2267, 1930, 2005, 2502, 1106, 2140, 1949, 1921,
2217, 2121, 2004, 1397, 1659, 1577, 2840, 1728, 1209, 1639

(a) Draw a frequency histogram to display these data. Use five classes. Is it reasonable to assume that the life span is normally distributed? Why?

(b) Find the mean and standard deviation of your sample.

(c) Compare the mean and standard deviation of your sample with those in the manufacturer's claim. Discuss the differences.

DATA **42.** ***Heights of Men*** You are performing a study about the height of 20- to 29-year-old men. A previous study found the height to be normally distributed, with a mean of 69.2 inches and a standard deviation of 2.9 inches. You randomly sample 30 men and find their heights to be as follows. *(Source: National Center for Health Statistics)*

72.1, 71.2, 67.9, 67.3, 69.5, 68.6, 68.8, 69.4, 73.5, 67.1,
69.2, 75.7, 71.1, 69.6, 70.7, 66.9, 71.4, 62.9, 69.2, 64.9,
68.2, 65.2, 69.7, 72.2, 67.5, 66.6, 66.5, 64.2, 65.4, 70.0

(a) Draw a frequency histogram to display these data. Use seven classes with midpoints of 63.85, 65.85, 67.85, 69.85, 71.85, 73.85, and 75.85. Is it reasonable to assume that the heights are normally distributed? Why?

(b) Find the mean and standard deviation of your sample.

(c) Compare the mean and standard deviation of your sample with those in the previous study. Discuss the differences.

Computing and Interpreting z-Scores of Normal Distributions In Exercises 43–46, you are given a normal distribution, the distribution's mean and standard deviation, four values from that distribution, and a graph of the Standard Normal Distribution. (a) Without converting to z-scores, match each value with the letters A, B, C, and D on the given graph of the Standard Normal Distribution. (b) Find the z-score that corresponds to each value and check your answers to part (a). (c) Determine whether any of the values are unusual.

43. ***Ball Bearings*** Your company manufactures ball bearings. The diameters of the ball bearings are normally distributed, with a mean of 3 inches and a standard deviation of 0.02 inch. The diameters of four ball bearings selected at random are 3.01, 2.97, 2.98, and 3.05.

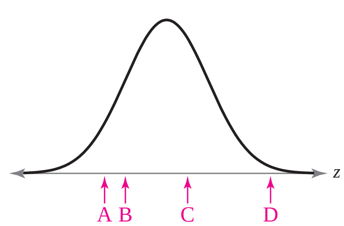


Figure for Exercise 43

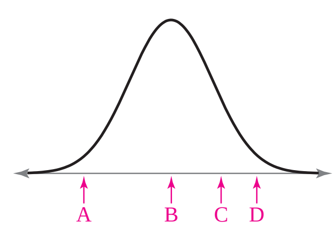


Figure for Exercise 44

43. (a) A = 2.97; B = 2.98; C = 3.01; D = 3.05
(b) 0.5; −1.5; −1; 2.5
(c) $x = 3.05$ is unusual owing to a relatively large z-score (2.5).

44. (a) A = 24,750; B = 30,000; C = 33,000; D = 35,150
(b) 2.06; −2.1; 0; 1.2
(c) $x = 35{,}150$ and $x = 24{,}750$ are unusual owing to their relatively large z-scores (2.06 and −2.1).

45. (a) A = 801; B = 950; C = 1250; D = 1467
(b) −0.36; 1.07; 2.11; −1.08
(c) $x = 1467$ is unusual owing to a relatively large z-score (2.11).

46. (a) A = 14; B = 18; C = 25; D = 32
(b) −0.58; 2.33; −1.42; 0.88
(c) $x = 32$ is unusual owing to a relatively large z-score (2.33).

47. 0.6915

48. 0.1587

49. 0.05

50. 0.8997

51. 0.5328

52. 0.2857

44. ***Tires*** An automobile tire brand has a life expectancy that is normally distributed, with a mean life of 30,000 miles and a standard deviation of 2500 miles. The life spans of four tires selected at random are 35,150 miles, 24,750 miles, 30,000 miles, and 33,000 miles.

45. ***SAT I Scores*** The SAT is an exam used by colleges and universities to evaluate undergraduate applicants. The test scores are normally distributed. In a recent year, the mean test score was 1026 and the standard deviation was 209. The test scores of four students selected at random are 950, 1250, 1467, and 801. *(Source: College Board Online)*

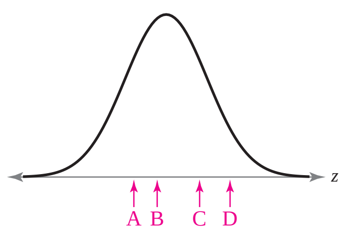


Figure for Exercise 45

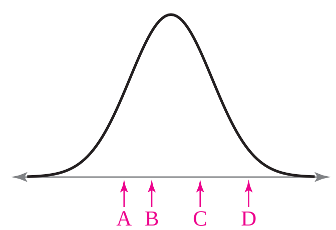
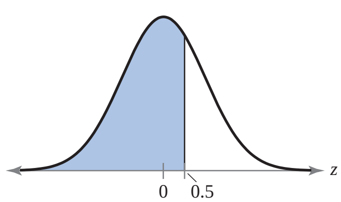


Figure for Exercise 46

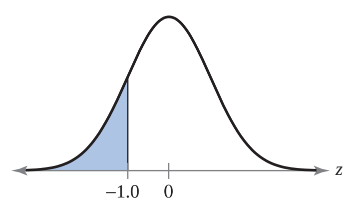
46. ***ACT Scores*** The ACT is an exam used by colleges and universities to evaluate undergraduate applicants. The test scores are normally distributed. In a recent year, the mean test score was 20.8 and the standard deviation was 4.8. The test scores of four students selected at random are 18, 32, 14, and 25. *(Source: ACT, Inc.)*

Graphical Analysis In Exercises 47–52, find the probability of z occurring in the indicated region.

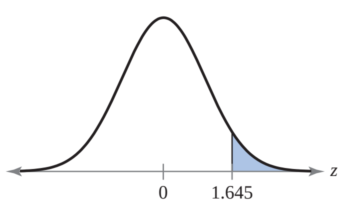
47.



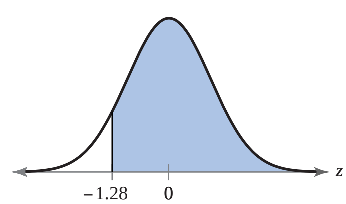
48.



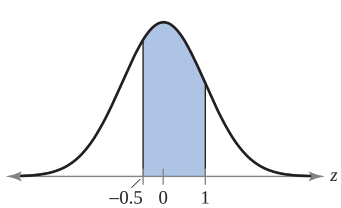
49.



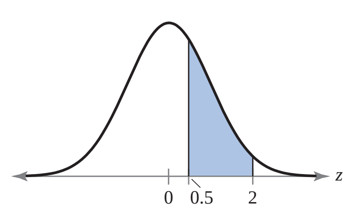
50.



51.

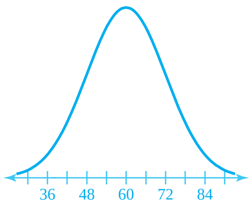


52.



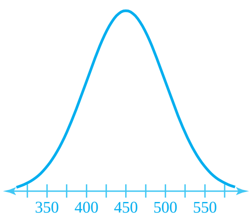
53. 0.9265 **54.** 0.6736
55. 0.9744 **56.** 0.5987
57. 0.3133 **58.** 0.4812
59. 0.901 **60.** 0.95
61. 0.0098 **62.** 0.05

63.



The normal distribution curve is centered at its mean (60) and has 2 points of inflection (48 and 72) representing $\mu \pm \sigma$.

64.



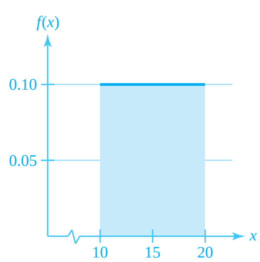
The normal distribution curve is centered at its mean (450) and has 2 points of inflection (400 and 500) representing $\mu \pm \sigma$.

65. (a) Area under curve
$= \text{area of rectangle}$
$= (1)(1)$
$= 1$

(b) 0.25

(c) 0.4

66. (a)



Area under curve
$= \text{area of rectangle}$
$= (20 - 10) \cdot (0.10)$
$= 1$

(b) 0.3

(c) 0.5

Finding Probabilities In Exercises 53–62, find the indicated probability using the standard normal distribution.

53. $P(z < 1.45)$ **54.** $P(z < 0.45)$ **55.** $P(z > -1.95)$

56. $P(z > -0.25)$ **57.** $P(-0.89 < z < 0)$ **58.** $P(-2.08 < z < 0)$

59. $P(-1.65 < z < 1.65)$ **60.** $P(-1.96 < z < 1.96)$

61. $P(z < -2.58 \text{ or } z > 2.58)$ **62.** $P(z < -1.96 \text{ or } z > 1.96)$

Extending Concepts

63. ***Writing*** Draw a normal curve with a mean of 60 and a standard deviation of 12. Describe how you constructed the curve and discuss its features.

64. ***Writing*** Draw a normal curve with a mean of 450 and a standard deviation of 50. Describe how you constructed the curve and discuss its features.

65. ***Uniform Distribution*** Another continuous distribution is the **uniform distribution.** An example is $f(x) = 1$ for $0 \le x \le 1$. The mean of this distribution for this example is 0.5 and the standard deviation is approximately 0.29. The graph of this distribution for this example is a square with the height and width both equal to 1 unit. In general, the density function for a uniform distribution on the interval from $x = a$ to $x = b$ is given by

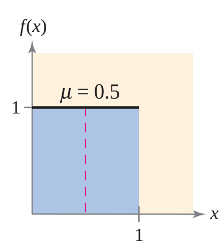
$$f(x) = \frac{1}{b - a}.$$

The mean is

$$\frac{a + b}{2}$$

and the variance is

$$\frac{(b - a)^2}{12}.$$



(a) Verify that the area under the curve is 1.

(b) Find the probability that x falls between 0.25 and 0.5.

(c) Find the probability that x falls between 0.3 and 0.7.

66. ***Uniform Distribution*** Consider the uniform density function $f(x) = 0.1$ for $10 \le x \le 20$. The mean of this distribution is 15 and the standard deviation is about 2.89.

(a) Draw a graph of the distribution and show that the area under the curve is 1.

(b) Find the probability that x falls between 12 and 15.

(c) Find the probability that x falls between 13 and 18.

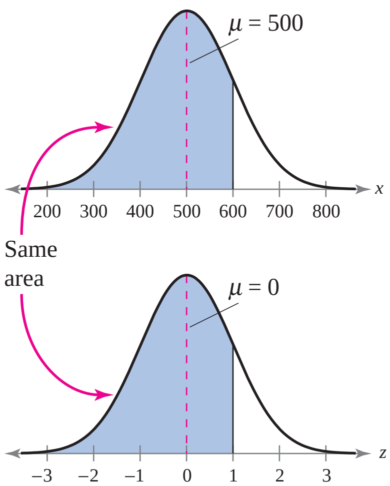
5.2 Normal Distributions: Finding Probabilities

What You Should Learn

- How to find probabilities for normally distributed variables using a table and using technology

Probability and Normal Distributions

Probability and Normal Distributions



If a random variable x is normally distributed, you can find the probability that x will fall in a given interval by calculating the area under the normal curve for the given interval. To find the area under any normal curve, first convert the upper and lower bounds of the interval to z-scores. Then use the standard normal distribution to find the area. For instance, consider a normal curve with $\mu = 500$ and $\sigma = 100$, as shown at the upper left. The value of x one standard deviation above the mean is $\mu + \sigma = 500 + 100 = 600$. Now consider the standard normal curve shown at the lower left. The value of z one standard deviation above the mean is $\mu + \sigma = 0 + 1 = 1$. Because a z-score of 1 corresponds to an x-value of 600, and areas are not changed with a transformation to a standard normal curve, the shaded areas in the graphs are equal.

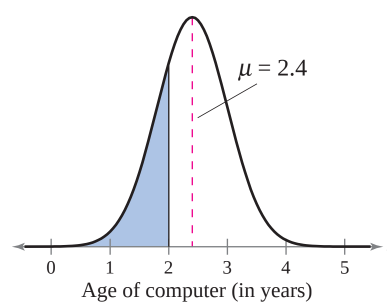
EXAMPLE 1

Finding Probabilities for Normal Distributions

A survey indicates that people use their computers an average of 2.4 years before upgrading to a new machine. The standard deviation is 0.5 year. A computer owner is selected at random. Find the probability that he or she will use it for less than 2 years before upgrading. Assume that the variable x is normally distributed.

SOLUTION The graph shows a normal curve with $\mu = 2.4$ and $\sigma = 0.5$ and a shaded area for x less than 2. The z-score that corresponds to 2 years is

$$z = \frac{x - \mu}{\sigma} = \frac{2 - 2.4}{0.5} = -0.80.$$



The Standard Normal Table shows that $P(z < -0.8) = 0.2119$. The probability that the computer will be upgraded in less than 2 years is 0.2119. So, 21.19% of new owners will upgrade in less than two years.

Try It Yourself 1

A Ford Focus manual transmission gets an average of 27 miles per gallon (mpg) in city driving with a standard deviation of 1.6 mpg. A Focus is selected at random. What is the probability that it will get more than 31 mpg? Assume that gas mileage is normally distributed. *(Source: U.S. Department of Energy)*

a. *Sketch* a graph.
b. *Find the z-score* that corresponds to 31 miles per gallon.
c. *Find the area* to the right of that z-score.
d. *Write* the result as a sentence.

Answer: Page A36

Study Tip

Another way to write the answer to Example 1 is $P(x < 2) = 0.2119$.

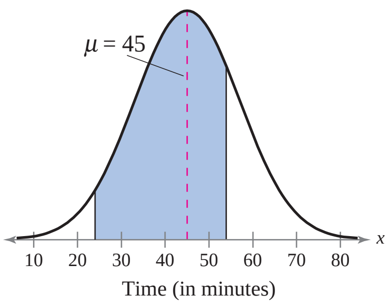
EXAMPLE 2

Finding Probabilities for Normal Distributions

A survey indicates that for each trip to the supermarket, a shopper spends an average of $\mu = 45$ minutes with a standard deviation of $\sigma = 12$ minutes. The length of time spent in the store is normally distributed and is represented by the variable x. A shopper enters the store. (a) Find the probability that the shopper will be in the store for each interval of time listed below. (b) If 200 shoppers enter the store, how many shoppers would you expect to be in the store for each interval of time listed below?

1. Between 24 and 54 minutes 2. More than 39 minutes

SOLUTION



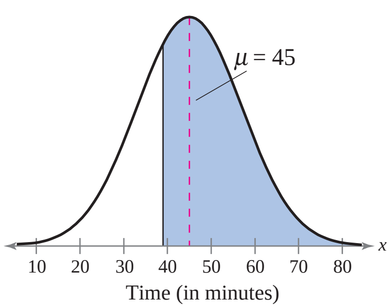
1.(a) The graph at the left shows a normal curve with $\mu = 45$ minutes and $\sigma = 12$ minutes. The area for x between 24 and 54 minutes is shaded. The z-scores that correspond to 24 minutes and to 54 minutes are

$$z_1 = \frac{24 - 45}{12} = -1.75 \quad \text{and} \quad z_2 = \frac{54 - 45}{12} = 0.75.$$

So, the probability that a shopper will be in the store between 24 and 54 minutes is

$$\begin{aligned} P(24 < x < 54) &= P(-1.75 < z < 0.75) \\ &= P(z < 0.75) - P(z < -1.75) \\ &= 0.7734 - 0.0401 = 0.7333. \end{aligned}$$

(b) Another way of interpreting this probability is to say that 73.33% of the shoppers will be in the store between 24 and 54 minutes. If 200 shoppers enter the store, then you would expect $200(0.7333) = 146.66$ (or about 147) shoppers to be in the store between 24 and 54 minutes.



2.(a) The graph at the left shows a normal curve with $\mu = 45$ minutes and $\sigma = 12$ minutes. The area for x greater than 39 minutes is shaded. The z-score that corresponds to 39 minutes is

$$z = \frac{39 - 45}{12} = -0.5.$$

So, the probability that a shopper will be in the store more than 39 minutes is

$$P(x > 39) = P(z > -0.5) = 1 - P(z < -0.5) = 1 - 0.3085 = 0.6915.$$

(b) If 200 shoppers enter the store, then you would expect $200(0.6915) = 138.3$ (or about 138) shoppers to be in the store more than 39 minutes.

Try It Yourself 2

What is the probability that the shopper will be in the supermarket between 33 and 60 minutes?

a. *Sketch* a graph.
b. *Find z-scores* that correspond to 60 minutes and 33 minutes.
c. *Find the cumulative area* for each z-score.
d. *Subtract the smaller area from the larger.*

Answer: Page A36

Another way to find normal probabilities is to use a calculator or a computer. You can find normal probabilities using MINITAB, Excel, and the TI-83.

EXAMPLE 3

Using Technology to Find Normal Probabilities

Assume that cholesterol levels of men in the United States are normally distributed, with a mean of 215 milligrams per deciliter and a standard deviation of 25 milligrams per deciliter. You randomly select a man from the United States. What is the probability that his cholesterol level is less than 175? Use a technology tool to find the probability.

SOLUTION MINITAB, Excel, and the TI-83 each have features that allow you to find normal probabilities without first converting to standard z-scores. For each, you must specify the mean and standard deviation of the population, as well as the x-value(s) that determine the interval.

MINITAB

```
Cumulative Distribution Function

Normal with mean = 215.000 and standard deviation = 25.0000

           x          P(X <= x)
    175.0000             0.0548
```

EXCEL

	A	B	C
1	NORMDIST(175,215,25,TRUE)		
2			0.054799

TI-83

```
normalcdf(0,175,215,25)
         .0547992894
```

From the displays, you can see that the probability that his cholesterol level is less than 175 is about 0.055, or 5.5%.

Try It Yourself 3

A man from the United States is selected at random. What is the probability that his cholesterol is between 190 and 225? Use a technology tool.

a. *Read the user's guide* for the technology tool you are using.
b. *Enter the appropriate data* to obtain the probability.
c. *Write* the result as a sentence.

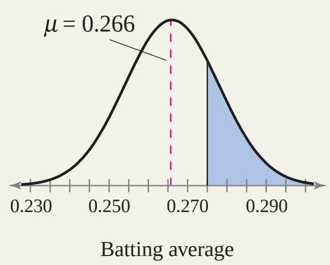
Answer: Page A36

Example 3 shows only one of several ways to find normal probabilities using MINITAB, Excel, and the TI-83.

Picturing the World

In baseball, a batting average is the number of hits divided by the number of at-bats. The batting averages of the more than 750 Major League Baseball players in a recent year can be approximated by a normal distribution, as shown in the following graph. The mean of the batting averages is 0.266 and the standard deviation is 0.012.

Major League Baseball



What percent of the players have a batting average of 0.275 or greater? If there are 40 players on a roster, how many would you expect to have a batting average of 0.275 or greater?

5.2 Exercises

Help

MyMathLab

Student Study Pack

1. 0.1151
2. 0.9974
3. 0.1151
4. 0.9861
5. 0.1144
6. 0.5434
7. 0.3022
8. 0.0878
9. 0.2742
10. 0.3462
11. 0.0566
12. 0.4251

Building Basic Skills and Vocabulary

Computing Probabilities In Exercises 1–6, assume the random variable x is normally distributed with mean $\mu = 86$ and standard deviation $\sigma = 5$. Find the indicated probability.

1. $P(x < 80)$

2. $P(x < 100)$

3. $P(x > 92)$

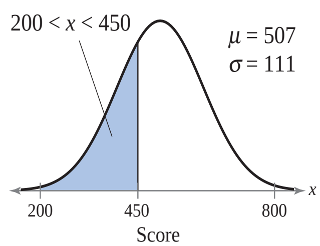
4. $P(x > 75)$

5. $P(70 < x < 80)$

6. $P(85 < x < 95)$

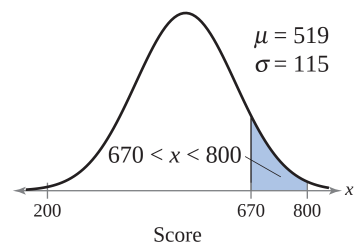
Graphical Analysis In Exercises 7–12, assume a member is selected at random from the population represented by the graph. Find the probability that the member selected at random is from the shaded area of the graph. Assume the variable x is normally distributed.

7. **SAT Verbal Scores**



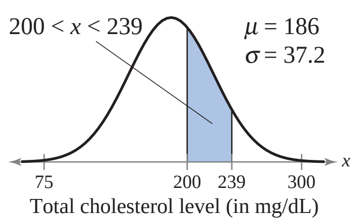
(Source: College Board Online)

8. **SAT Math Scores**



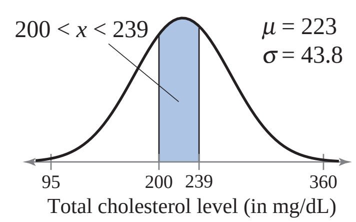
(Source: College Board Online)

9. **U.S. Women Ages 20–34: Total Cholesterol**



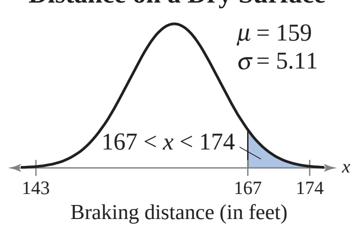
(Adapted from Centers for Disease Control and Prevention)

10. **U.S. Women Ages 55–64: Total Cholesterol**



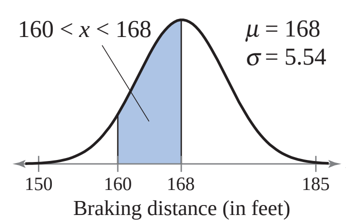
(Adapted from Centers for Disease Control and Prevention)

11. **Chevrolet Blazer: Braking Distance on a Dry Surface**



(Source: National Highway Traffic Safety Administration)

12. **Chevrolet Blazer: Braking Distance on a Wet Surface**



(Source: National Highway Traffic Safety Administration)

13. (a) 0.1357
(b) 0.6983
(c) 0.1660
14. (a) 0.0668
(b) 0.927
(c) 0.0062
15. (a) 0.1711
(b) 0.7018
(c) 0.1271
16. (a) 0.2514
(b) 0.4972
(c) 0.2514
17. (a) 0.0062
(b) 0.9876
(c) 0.0062

Using and Interpreting Concepts

Finding Probabilities In Exercises 13–20, find the indicated probabilities. If convenient, use technology to find the probabilities.

13. ***Heights of Men*** A survey was conducted to measure the height of U.S. men. In the survey, respondents were grouped by age. In the 20–29 age group, the heights were normally distributed, with a mean of 69.2 inches and a standard deviation of 2.9 inches. A study participant is randomly selected. *(Source: U.S. National Center for Health Statistics)*

(a) Find the probability that his height is less than 66 inches.
(b) Find the probability that his height is between 66 and 72 inches.
(c) Find the probability that his height is more than 72 inches.

14. ***Fish Lengths*** The lengths of Atlantic croaker fish are normally distributed, with a mean of 10 inches and a standard deviation of 2 inches. An Atlantic croaker fish is randomly selected. *(Adapted from National Marine Fisheries Service, Fisheries Statistics and Economics Division)*

(a) Find the probability that the length of the fish is less than 7 inches.
(b) Find the probability that the length of the fish is between 7 and 15 inches.
(c) Find the probability that the length of the fish is more than 15 inches.

15. ***ACT Scores*** In a recent year, the ACT scores for high school students with a 3.50 to 4.00 grade point average were normally distributed, with a mean of 24.1 and a standard deviation of 4.3. A student with a 3.50 to 4.00 grade point average who took the ACT during this time is randomly selected. *(Source: ACT, Inc.)*

(a) Find the probability that the student's ACT score is less than 20.
(b) Find the probability that the student's ACT score is between 20 and 29.
(c) Find the probability that the student's ACT score is more than 29.

16. ***Rhesus Monkeys*** The weights of adult male rhesus monkeys are normally distributed, with a mean of 15 pounds and a standard deviation of 3 pounds. A rhesus monkey is randomly selected.

(a) Find the probability that the monkey's weight is less than 13 pounds.
(b) Find the probability that the weight is between 13 and 17 pounds.
(c) Find the probability that the monkey's weight is more than 17 pounds.

17. ***Computer Usage*** A survey was conducted to measure the number of hours per week adults in the United States spend on home computers. In the survey, the number of hours were normally distributed, with a mean of 7 hours and a standard deviation of 1 hour. A survey participant is randomly selected.

(a) Find the probability that the hours spent on the home computer by the participant are less than 4.5 hours per week.
(b) Find the probability that the hours spent on the home computer by the participant are between 4.5 and 9.5 hours per week.
(c) Find the probability that the hours spent on the home computer by the participant are more than 9.5 hours per week.

18. (a) 0.0475
(b) 0.8469
(c) 0.1056
19. (a) 0.0073
(b) 0.806
(c) 0.1867
20. (a) 0.2743
(b) 0.3811
(c) 0.3446
21. (a) 79.95%
(b) 348
22. (a) 43.25%
(b) 363
23. (a) 64.8%
(b) 18

18. ***Utility Bills*** The monthly utility bills in a city are normally distributed, with a mean of \$100 and a standard deviation of \$12. A utility bill is randomly selected.

(a) Find the probability that the utility bill is less than \$80.
(b) Find the probability that the utility bill is between \$80 and \$115.
(c) Find the probability that the utility bill is more than \$115.

19. ***Computer Lab Schedule*** The time per week a student uses a lab computer is normally distributed, with a mean of 6.2 hours and a standard deviation of 0.9 hour. A student is randomly selected.

(a) Find the probability that the student uses a lab computer less than 4 hours per week.
(b) Find the probability that the student uses a lab computer between 4 and 7 hours per week.
(c) Find the probability that the student uses a lab computer more than 7 hours per week.

20. ***Health Club Schedule*** The time per workout an athlete uses a stairclimber is normally distributed, with a mean of 20 minutes and a standard deviation of 5 minutes. An athlete is randomly selected.

(a) Find the probability that the athlete uses a stairclimber for less than 17 minutes.
(b) Find the probability that the athlete uses a stairclimber between 17 and 22 minutes.
(c) Find the probability that the athlete uses a stairclimber for more than 22 minutes.

Using Normal Distributions In Exercises 21–30, answer the questions about the specified normal distribution.

21. ***SAT Verbal Scores*** Use the normal distribution of SAT verbal scores in Exercise 7 for which the mean is 507 and the standard deviation is 111.

(a) What percent of the SAT verbal scores are less than 600?
(b) If 1000 SAT verbal scores are randomly selected, about how many would you expect to be greater than 550?

22. ***SAT Math Scores*** Use the normal distribution of SAT math scores in Exercise 8 for which the mean is 519 and the standard deviation is 115.

(a) What percent of the SAT math scores are less than 500?
(b) If 1500 SAT math scores are randomly selected, about how many would you expect to be greater than 600?

23. ***Cholesterol*** Use the normal distribution of women's total cholesterol levels in Exercise 9 for which the mean is 186 milligrams per deciliter and the standard deviation is 37.2 milligrams per deciliter.

(a) What percent of the women have a total cholesterol level less than 200 milligrams per deciliter of blood?
(b) If 250 U.S. women in the 20–29 age group are randomly selected, about how many would you expect to have a total cholesterol level greater than 240 milligrams per deciliter of blood?

24. (a) 0.6443
(b) 140
25. (a) 30.85%
(b) 31
26. (a) 4.75%
(b) 7
27. (a) 99.87%
(b) 0.798
28. (a) 1.88%
(b) 60
29. 1.5%; It is unusual for a battery to have a life span that is more than 2065 hours because of the relatively large *z*-score (2.17).
30. 5.94%; It is not unusual for a person to consume less than 3.1 pounds of peanuts in a year because the *z*-score is within 2 standard deviations of the mean.

24. *Cholesterol* Use the normal distribution of women's total cholesterol levels in Exercise 10 for which the mean is 223 milligrams per deciliter and the standard deviation is 43.8 milligrams per deciliter.

(a) What percent of the women have a total cholesterol level less than 239 milligrams per deciliter of blood?

(b) If 200 U.S. women in the 50–59 age group are randomly selected, about how many would you expect to have a total cholesterol level greater than 200 milligrams per deciliter of blood?

25. *Fish Lengths* Use the normal distribution of fish lengths in Exercise 14 for which the mean is 10 inches and the standard deviation is 2 inches.

(a) What percent of the fish are longer than 11 inches?

(b) If 200 Atlantic croakers are randomly selected, about how many would you expect to be shorter than 8 inches?

26. *Rhesus Monkeys* Use the normal distribution of monkey weights in Exercise 16 for which the mean is 15 pounds and the standard deviation is 3 pounds.

(a) What percent of the monkeys have a weight that is greater than 20 pounds?

(b) If 50 rhesus monkeys are randomly selected, about how many would you expect to weigh less than 12 pounds?

27. *Computer Usage* Use the normal distribution of computer usage in Exercise 17 for which the mean is 7 hours and the standard deviation is 1 hour.

(a) What percent of the adults spend more than 4 hours per week on a home computer?

(b) If 35 adults in the United States are randomly selected, about how many would you expect to say they spend less than 5 hours per week on a home computer?

28. *Utility Bills* Use the normal distribution of utility bills in Exercise 18 for which the mean is $100 and the standard deviation is $12.

(a) What percent of the utility bills are more than $125?

(b) If 300 utility bills are randomly selected, about how many would you expect to be less than $90?

29. *Battery Life Spans* The life span of a battery is normally distributed, with a mean of 2000 hours and a standard deviation of 30 hours. What percent of batteries have a life span that is more than 2065 hours? Would it be unusual for a battery to have a life span that is more than 2065 hours? Explain your reasoning.

30. *Peanuts* Assume the mean annual consumption of peanuts is normally distributed, with a mean of 5.9 pounds per person and a standard deviation of 1.8 pounds per person. What percent of people annually consume less than 3.1 pounds of peanuts per person? Would it be unusual for a person to consume less than 3.1 pounds of peanuts in a year? Explain your reasoning.

31. Out of control, because there is a point more than 3 standard deviations beyond the mean.

32. Out of control, because two out of three consecutive points lie more than 2 standard deviations from the mean.

33. Out of control, because there are nine consecutive points below the mean, and two out of three consecutive points lie more than 2 standard deviations from the mean.

34. In control, because none of the three warning signals detected a change.

Extending Concepts

Control Charts **Statistical process control (SPC)** is the use of statistics to monitor and improve the quality of a process, such as manufacturing an engine part. In SPC, information about a process is gathered and used to determine if a process is meeting all of the specified requirements. One tool used in SPC is a **control chart.** When individual measurements of a variable x are normally distributed, a control chart can be used to detect processes that are possibly out of statistical control. Three warning signals that a control chart uses to detect a process that may be out of control are as follows:

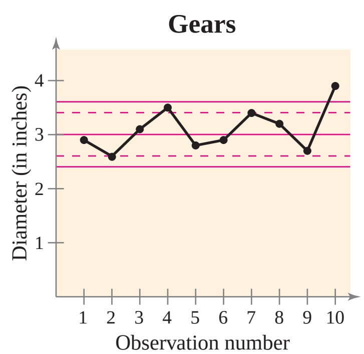
(1) A point lies beyond three standard deviations of the mean.

(2) There are nine consecutive points that fall on one side of the mean.

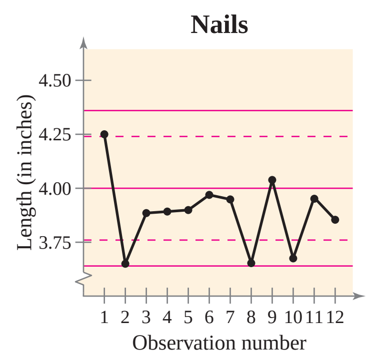
(3) At least two of three consecutive points lie more than two standard deviations from the mean.

In Exercises 31–34, a control chart is shown. Each chart has horizontal lines drawn at the mean μ, at $\mu \pm 2\sigma$, and at $\mu \pm 3\sigma$. Determine if the process shown is in control or out of control. Explain.

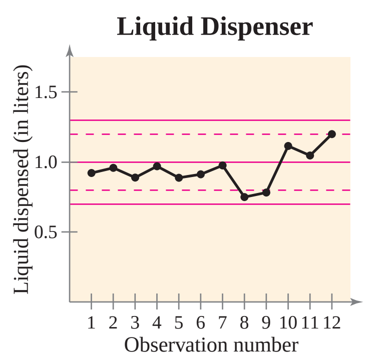
31. A gear has been designed to have a diameter of 3 inches. The standard deviation of the process is 0.2 inch.



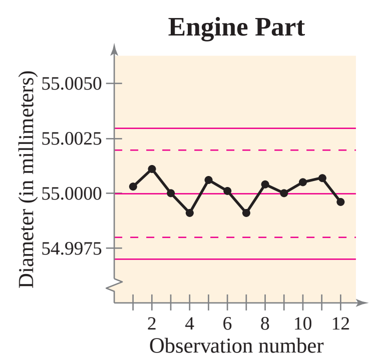
32. A nail has been designed to have a length of 4 inches. The standard deviation of the process is 0.12 inch.



33. A liquid-dispensing machine has been designed to fill bottles with 1 liter of liquid. The standard deviation of the process is 0.1 liter.



34. An engine part has been designed to have a diameter of 55 millimeters. The standard deviation of the process is 0.001 millimeter.



5.3 Normal Distributions: Finding Values

What You Should Learn

- How to find a z-score given the area under the normal curve
- How to transform a z-score to an x-value
- How to find a specific data value of a normal distribution given the probability

Finding z-Scores • Transforming a z-Score to an x-Value • Finding a Specific Data Value for a Given Probability

Finding z-Scores

In Section 5.2, you were given a normally distributed random variable x and you found the probability that x would fall in a given interval by calculating the area under the normal curve for the given interval.

But what if you are given a probability and want to find a value? For instance, a university might want to know what is the lowest test score a student can have on an entrance exam and still be in the top 10%, or a medical researcher might want to know the cutoff values to select the middle 90% of patients by age. In this section, you will learn how to find a value given an area under a normal curve (or a probability), as shown in the following example.

Note to Instructor

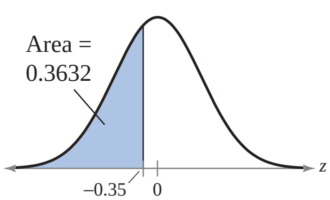
Have students note that as the z-scores increase, the cumulative areas increase. The CDF is one to one and as such has an inverse function. Discuss how this inverse function (INVCDF) can be used when a cumulative area (percentile) is known and the z-score must be found.

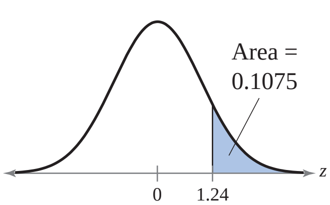
Study Tip

You can use a computer or calculator to find the z-score that corresponds to a cumulative area. For instance, here are instructions for finding the z-score that corresponds to an area of 0.3632 on a TI-83.

2nd DISTR 3 .3632

The calculator will display −.3499183227.





EXAMPLE 1

Finding a z-Score Given an Area

1. Find the z-score that corresponds to a cumulative area of 0.3632.
2. Find the z-score that has 10.75% of the distribution's area to its right.

SOLUTION

1. Find the z-score that corresponds to an area of 0.3632 by locating 0.3632 in the Standard Normal Table. The values at the beginning of the corresponding row and at the top of the corresponding column give the z-score. For this area, the row value is −0.3 and the column value is 0.05. So, the z-score is −0.35.

z	.09	.08	.07	.06	.05	.04	.03
−3.4	.0002	.0003	.0003	.0003	.0003	.0003	.0003
−0.5	.2776	.2810	.2843	.2877	.2912	.2946	.2981
−0.4	.3121	.3156	.3192	.3228	.3264	.3300	.3336
−0.3	.3483	.3520	.3557	.3594	.3632	.3669	.3707
−0.2	.3859	.3897	.3936	.3974	.4013	.4052	.4090

2. Because the area to the right is 0.1075, the cumulative area is $1 - 0.1075 = 0.8925$. Find the z-score that corresponds to an area of 0.8925 by locating 0.8925 in the Standard Normal Table. For this area, the row value is 1.2 and the column value is 0.04. So, the z-score is 1.24.

z	.00	.01	.02	.03	.04	.05	.06
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131

Note to Instructor

If you prefer that your students use a 0-to-z table, refer them to Appendix A where an alternative presentation for this material is given.

Study Tip

In most cases, the given area will not be found in the table, so use the entry closest to it. If the given area is halfway between two area entries, use the z-score halfway between the corresponding z-scores. For instance, in part 1 of Example 2, the z-score between -1.64 and -1.65 is -1.645.

Try It Yourself 1

1. Find the z-score that has 96.16% of the distribution's area to the right.
2. Find the z-score for which 95% of the distribution's area lies between $-z$ and z.

a. *Determine* the cumulative area.
b. *Locate* the area in the Standard Normal Table.
c. *Find the z-score* that corresponds to the area.

Answer: Page A36

In Section 2.5, you learned that percentiles divide a data set into one hundred equal parts. To find a z-score that corresponds to a percentile, you can use the Standard Normal Table. Recall that if a value x represents the 83rd percentile P_{83}, then 83% of the data values are below x and 17% of the data values are above x.

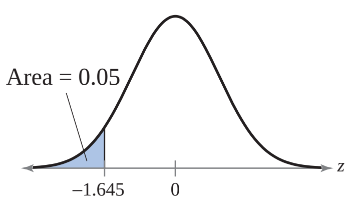
EXAMPLE 2

Finding a z-Score Given a Percentile

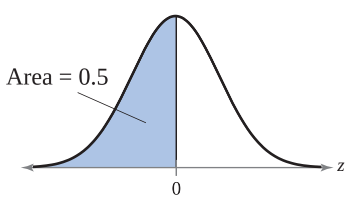
Find the z-score that corresponds to each percentile.

1. P_5 2. P_{50} 3. P_{90}

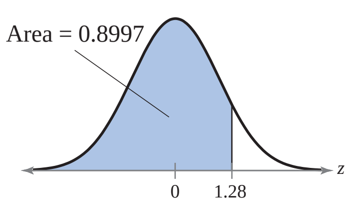
SOLUTION



1. To find the z-score that corresponds to P_5, find the z-score that corresponds to an area of 0.05 (see figure) by locating 0.05 in the Standard Normal Table. The areas closest to 0.05 in the table are 0.0495 $(z = -1.65)$ and 0.0505 $(z = -1.64)$. Because 0.05 is halfway between the two areas in the table, use the z-score that is halfway between -1.64 and -1.65. So, the z-score that corresponds to an area of 0.05 is -1.645.



2. To find the z-score that corresponds to P_{50}, find the z-score that corresponds to an area of 0.5 (see figure) by locating 0.5 in the Standard Normal Table. The area closest to 0.5 in the table is 0.5000, so the z-score that corresponds to an area of 0.5 is 0.00.



3. To find the z-score that corresponds to P_{90}, find the z-score that corresponds to an area of 0.9 (see figure) by locating 0.9 in the Standard Normal Table. The area closest to 0.9 in the table is 0.8997, so the z-score that corresponds to an area of 0.9 is 1.28.

Try It Yourself 2

Find the z-score that corresponds to each percentile.

1. P_{10} 2. P_{20} 3. P_{99}

a. *Write* the percentile as an area. If necessary, draw a graph of the area to visualize the problem.
b. *Locate* the area in the Standard Normal Table. If the area is not in the table, use the closest area. (See Study Tip above.)
c. *Identify* the z-score that corresponds to the area.

Answer: Page A36

Transforming a z-Score to an x-Value

Recall that to transform an x-value to a z-score, you can use the formula

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

This formula gives z in terms of x. If you solve this formula for x, you get a new formula that gives x in terms of z.

$$z = \frac{x - \mu}{\sigma} \qquad \text{Formula for } z \text{ in terms of } x$$

$$z\sigma = x - \mu \qquad \text{Multiply each side by } \sigma.$$

$$\mu + z\sigma = x \qquad \text{Add } \mu \text{ to each side.}$$

$$x = \mu + z\sigma \qquad \text{Interchange sides.}$$

Transforming a z-Score to an x-Value

To transform a standard z-score to a data value x in a given population, use the formula

$$x = \mu + z\sigma.$$

EXAMPLE 3

Finding an x-Value Corresponding to a z-Score

The speeds of vehicles along a stretch of highway are normally distributed, with a mean of 56 miles per hour and a standard deviation of 4 miles per hour. Find the speeds x corresponding to z-scores of 1.96, -2.33, and 0. Interpret your results.

SOLUTION The x-value that corresponds to each standard score is calculated using the formula $x = \mu + z\sigma$.

$z = 1.96$: $\quad x = 56 + 1.96(4) = 63.84$ miles per hour

$z = -2.33$: $\quad x = 56 + (-2.33)(4) = 46.68$ miles per hour

$z = 0$: $\quad x = 56 + 0(4) = 56$ miles per hour

Interpretation You can see that 63.84 miles per hour is above the mean, 46.68 is below the mean, and 56 is equal to the mean.

Try It Yourself 3

The monthly utility bills in a city are normally distributed, with a mean of \$70 and a standard deviation of \$8. Find the x-values that correspond to z-scores of -0.75, 4.29, and -1.82. What can you conclude?

a. *Identify* μ and σ of the nonstandard normal distribution.
b. *Transform* each z-score to an x-value.
c. *Interpret* the results.

Answer: Page A36

Finding a Specific Data Value for a Given Probability

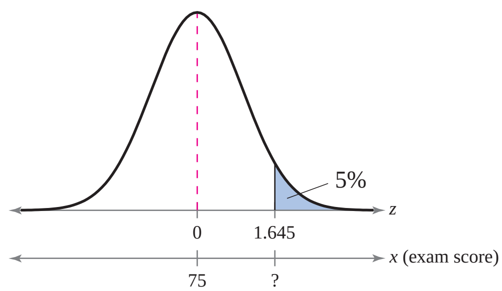
You can also use the normal distribution to find a specific data value (x-value) for a given probability, as shown in Example 4.

EXAMPLE 4

Finding a Specific Data Value

Scores for a civil service exam are normally distributed, with a mean of 75 and a standard deviation of 6.5. To be eligible for civil service employment, you must score in the top 5%. What is the lowest score you can earn and still be eligible for employment?

SOLUTION Exam scores in the top 5% correspond to the shaded region shown.



An exam score in the top 5% is any score above the 95th percentile. To find the score that represents the 95th percentile, you must first find the z-score that corresponds to a cumulative area of 0.95. From the Standard Normal Table, you can find that the areas closest to 0.95 are 0.9495 $(z = 1.64)$ and 0.9505 $(z = 1.65)$. Because 0.95 is halfway between the two areas in the table, use the z-score that is halfway between 1.64 and 1.65. That is, $z = 1.645$. Using the equation $x = \mu + z\sigma$, you have

$$\begin{aligned} x &= \mu + z\sigma \\ &= 75 + 1.645(6.5) \\ &\approx 85.69. \end{aligned}$$

Interpretation The lowest score you can earn and still be eligible for employment is 86.

Try It Yourself 4

The braking distances of a sample of Ford F-150s are normally distributed. On a dry surface, the mean braking distance was 158 feet and the standard deviation was 6.51 feet. What is the longest braking distance on a dry surface one of these Ford F-150s could have and still be in the top 1%? *(Adapted from National Highway Traffic Safety Administration)*

a. *Sketch* a graph.
b. *Find the z-score* that corresponds to the given area.
c. *Find x* using the equation $x = \mu + z\sigma$.
d. *Interpret* the result.

Answer: Page A36

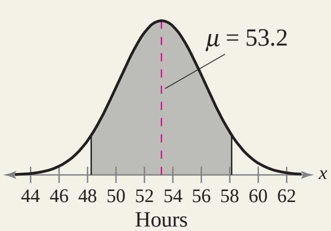
Note to Instructor

Mention that to use the cumulative table to find a z-score, you must first express the area given as a cumulative area. It helps to explain these as percentiles. For example, the score in the top 20% represents the 80th percentile. Point out that if students are using a technology tool to find an x-value that corresponds to an area, it is not necessary first to find a z-score.

Picturing the World

According to the American Medical Association, the mean number of hours all physicians spend in patient care each week is about 53.2 hours. The hours spent in patient care each week by physicians can be approximated by a normal distribution. Assume the standard deviation is 3 hours.

Hours Physicians Spend in Patient Care



Between what two values does the middle 90% of the data lie?

EXAMPLE 5

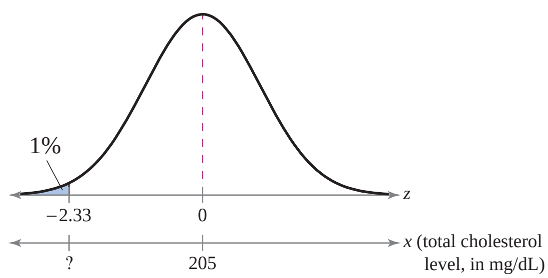
Finding a Specific Data Value

In a randomly selected sample of 1169 men ages 35–44, the mean total cholesterol level was 205 milligrams per deciliter with a standard deviation of 39.2 milligrams per deciliter. Assume the total cholesterol levels are normally distributed. Find the highest total cholesterol level a man in this 35–44 age group can have and be in the lowest 1%. *(Adapted from Centers for Disease Control and Prevention)*

SOLUTION

Total cholesterol levels in the lowest 1% correspond to the shaded region shown.

Total Cholesterol Levels in Men Ages 35–44



A total cholesterol level in the lowest 1% is any level below the 1st percentile. To find the level that represents the 1st percentile, you must first find the z-score that corresponds to a cumulative area of 0.01. From the Standard Normal Table, you can find that the area closest to 0.01 is 0.0099. So, the z-score that corresponds to an area of 0.01 is $z = -2.33$. Using the equation $x = \mu + z\sigma$, you have

$$\begin{aligned} x &= \mu + z\sigma \\ &= 205 + (-2.33)(39.2) \\ &\approx 113.66. \end{aligned}$$

Interpretation The value that separates the lowest 1% of total cholesterol levels for men in the 35–44 age group from the highest 99% is about 114.

Try It Yourself 5

The length of time employees have worked at a corporation is normally distributed, with a mean of 11.2 years and a standard deviation of 2.1 years. In a company cutback, the lowest 10% in seniority are laid off. What is the maximum length of time an employee could have worked and still be laid off?

a. *Sketch* a graph.
b. *Find the z-score* that corresponds to the given area.
c. *Find x* using the equation $x = \mu + z\sigma$.
d. *Interpret* the result.

Answer: Page A36

5.3 Exercises

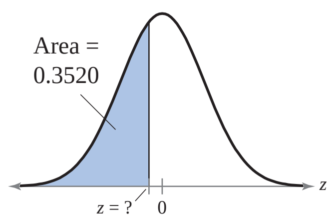
Building Basic Skills and Vocabulary

In Exercises 1–24, use the Standard Normal Table to find the z-score that corresponds to the given cumulative area or percentile. If the area is not in the table, use the entry closest to the area. If the area is halfway between two entries, use the z-score halfway between the corresponding z-scores.

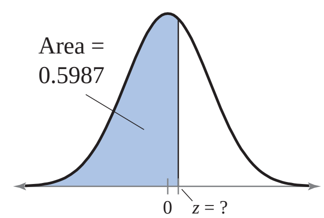
1. 0.0202	**2.** 0.2090	**3.** 0.8023	**4.** 0.6443
5. 0.4364	**6.** 0.0080	**7.** 0.9916	**8.** 0.7673
9. 0.05	**10.** 0.85	**11.** 0.84	**12.** 0.01
13. P_1	**14.** P_{15}	**15.** P_{40}	**16.** P_{55}
17. P_{88}	**18.** P_{96}	**19.** P_{25}	**20.** P_{50}
21. P_{75}	**22.** P_{10}	**23.** P_{35}	**24.** P_{65}

Graphical Analysis In Exercises 25–30, find the indicated z-score(s) shown in the graph.

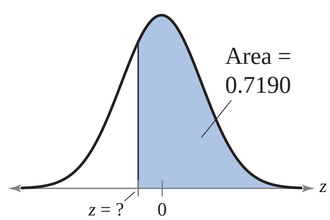
25.



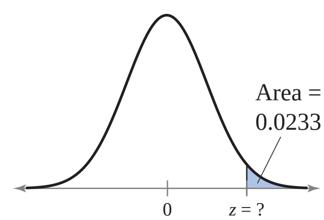
26.



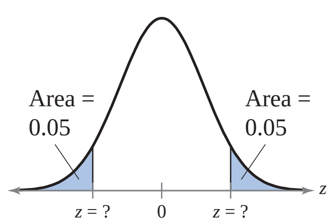
27.



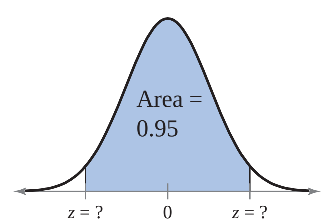
28.



29.



30.



In Exercises 31–38, find the indicated z-score.

31. Find the z-score that has 62.8% of the distribution's area to its left.

32. Find the z-score that has 78.5% of the distribution's area to its left.

33. Find the z-score that has 62.8% of the distribution's area to its right.

34. Find the z-score that has 78.5% of the distribution's area to its right.

35. Find the z-score for which 80% of the distribution's area lies between $-z$ and z.

1. -2.05
2. -0.81
3. 0.85
4. 0.37
5. -0.16
6. -2.41
7. 2.39
8. 0.73
9. -1.645
10. 1.04
11. 0.995
12. -2.325
13. -2.325
14. -1.04
15. -0.25
16. 0.125
17. 1.175
18. 1.75
19. -0.675
20. 0
21. 0.675
22. -1.28
23. -0.385
24. 0.385
25. -0.38
26. 0.25
27. -0.58
28. 1.99
29. -1.645, 1.645
30. ± 1.96
31. 0.325

32. 0.79
33. −0.33
34. −0.79
35. 1.28
36. 2.575
37. ±0.06
38. ±0.15
39. (a) 68.52
(b) 62.14
40. (a) 72.91
(b) 67.24
41. (a) 12.28
(b) 20.08
42. (a) 6.765
(b) 13.725

36. Find the z-score for which 99% of the distribution's area lies between $-z$ and z.

37. Find the z-score for which 5% of the distribution's area lies between $-z$ and z.

38. Find the z-score for which 12% of the distribution's area lies between $-z$ and z.

Using and Interpreting Concepts

Using Normal Distributions In Exercises 39–44, answer the questions about the specified normal distribution.

39. ***Heights of Women*** In a survey of women in the United States (ages 20–29), the mean height was 64 inches with a standard deviation of 2.75 inches. *(Source: National Center for Health Statistics)*

(a) What height represents the 95th percentile?
(b) What height represents the first quartile?

40. ***Heights of Men*** In a survey of men in the United States (ages 20–29), the mean height was 69.2 inches with a standard deviation of 2.9 inches. *(Source: National Center for Health Statistics)*

(a) What height represents the 90th percentile?
(b) What height represents the first quartile?

41. ***Apples*** The annual per capita utilization of apples (in pounds) in the United States can be approximated by a normal distribution, as shown in the graph. *(Adapted from U.S. Department of Agriculture)*

(a) What annual per capita utilization of apples represents the 10th percentile?
(b) What annual per capita utilization of apples represents the third quartile?

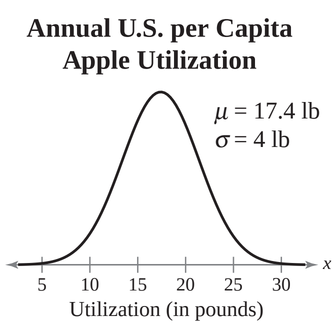


Figure for Exercise 41

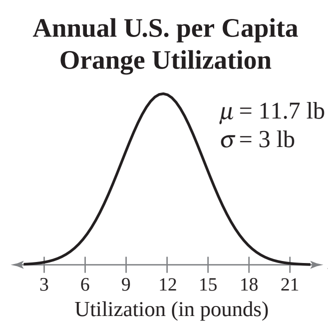


Figure for Exercise 42

42. ***Oranges*** The annual per capita utilization of oranges (in pounds) in the United States can be approximated by a normal distribution, as shown in the graph. *(Adapted from U.S. Department of Agriculture)*

(a) What annual per capita utilization of oranges represents the 5th percentile?
(b) What annual per capita utilization of oranges represents the third quartile?

Time Spent Waiting for a Heart

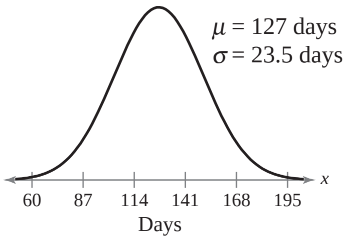


Figure for Exercise 43

Annual U.S. per Capita Ice Cream Consumption

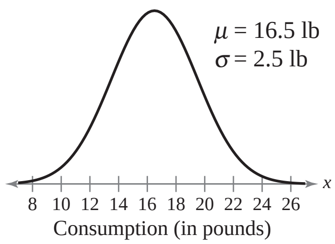


Figure for Exercise 44

43. (a) 139.22
(b) 96.92
44. (a) 18.1875
(b) 13.9125
45. 19.89
46. 15.1224
47. Tires that wear out by 26,800 miles will be replaced free of charge.
48. A = 83.52; B = 76.68; C = 67.32; D = 60.48
49. 7.93

Final Exam Grades

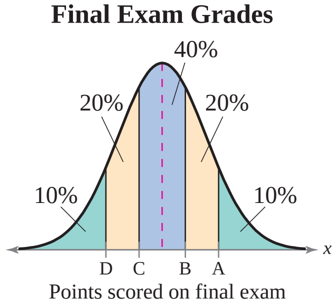


Figure for Exercise 48

43. ***Heart Transplant Waiting Times*** The time spent (in days) waiting for a heart transplant in Ohio and Michigan for patients with type A^+ blood can be approximated by a normal distribution, as shown in the graph. *(Source: Organ Procurement and Transplant Network)*

(a) What is the shortest time spent waiting for a heart that would still place a patient in the top 30% of waiting times?

(b) What is the longest time spent waiting for a heart that would still place a patient in the bottom 10% of waiting times?

44. ***Ice Cream*** The annual per capita consumption of ice cream (in pounds) in the United States can be approximated by a normal distribution, as shown in the graph. *(Adapted from U.S. Department of Agriculture)*

(a) What is the smallest annual per capita consumption of ice cream that can be in the top 25% of consumptions?

(b) What is the largest annual per capita consumption of ice cream that can be in the bottom 15% of consumptions?

45. ***Cereal Boxes*** The weights of the contents of a cereal box are normally distributed with a mean weight of 20 ounces and a standard deviation of 0.07 ounce. Boxes in the lower 5% do not meet the minimum weight requirements and must be repackaged. What is the minimum weight requirement for a cereal box?

46. ***Bags of Cookies*** The weights of bags of cookies are normally distributed with a mean of 15 ounces and a standard deviation of 0.085 ounce. Bags of cookies that have weights in the upper 7.5% are too heavy and must be repackaged. What is the most a bag of cookies can weigh and not need to be repackaged?

Extending Concepts

47. ***Writing a Guarantee*** You sell a brand of automobile tire that has a life expectancy that is normally distributed, with a mean life of 30,000 miles and a standard deviation of 2500 miles. You want to give a guarantee for free replacement of tires that don't wear well. How should you word your guarantee if you are willing to replace approximately 10% of the tires you sell?

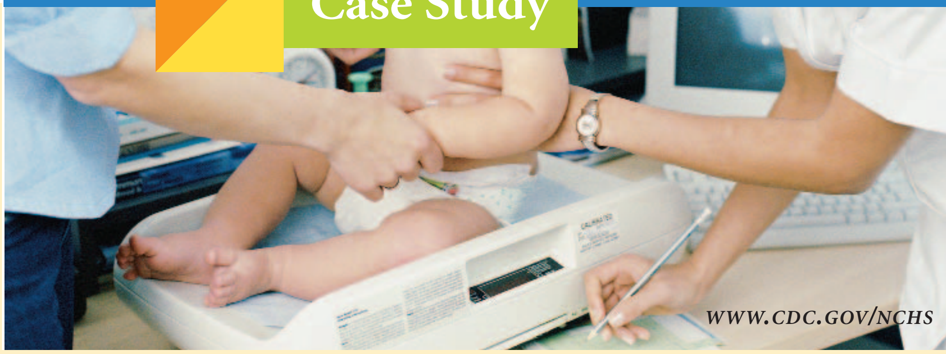
48. ***Statistics Grades*** In a large section of a statistics class, the points for the final exam are normally distributed with a mean of 72 and a standard deviation of 9. Grades are to be assigned according to the following rule.

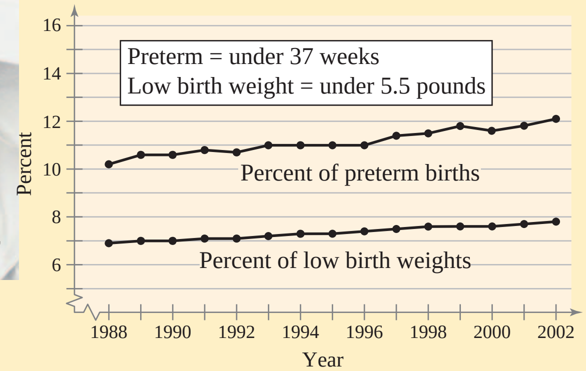
- The top 10% receive As
- The next 20% receive Bs
- The middle 40% receive Cs
- The next 20% receive Ds
- The bottom 10% receive Fs

Find the lowest score on the final exam that would qualify a student for an A, a B, a C, and a D.

49. ***Vending Machine*** A vending machine dispenses coffee into an eight-ounce cup. The amount of coffee dispensed into the cup is normally distributed with a standard deviation of 0.03 ounce. You can allow the cup to overfill 1% of the time. What amount should you set as the mean amount of coffee to be dispensed?

Case Study





Birth Weights in America

The National Center for Health Statistics (NCHS) keeps records of many health-related aspects of people, including the birth weights of all babies born in the United States.

The birth weight of a baby is related to its gestation period (the time between conception and birth). For a given gestation period, the birth weights can be approximated by a normal distribution. The means and standard deviations of the birth weights for various gestation periods are shown at the right.

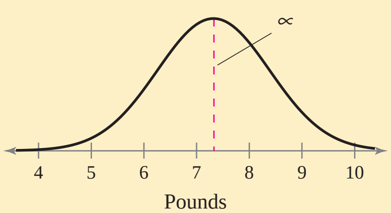
One of the many goals of the NCHS is to reduce the percentage of babies born with low birth weights. As you can see from the graph at the upper right, the problem of low birth weights increased from 1988 to 2002.

Gestation period	Mean birth weight	Standard deviation
Under 28 Weeks	1.88 lb	1.19 lb
28 to 31 Weeks	4.07 lb	1.87 lb
32 to 35 Weeks	5.73 lb	1.48 lb
36 Weeks	6.46 lb	1.20 lb
37 to 39 Weeks	7.33 lb	1.09 lb
40 Weeks	7.72 lb	1.05 lb
41 Weeks	7.83 lb	1.08 lb
42 Weeks and over	7.65 lb	1.12 lb

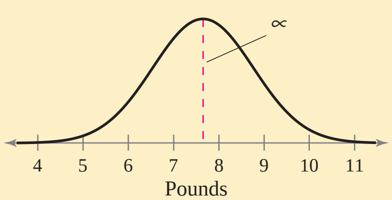
Exercises

1. The distributions of birth weights for three gestation periods are shown. Match the curves with the gestation periods. Explain your reasoning.

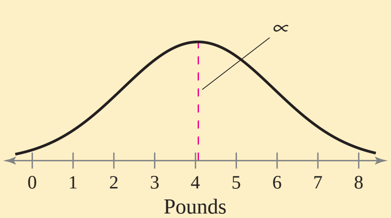
(a)



(b)



(c)



2. What percent of the babies born with each gestation period have a low birth weight (under 5.5 pounds)? Explain your reasoning.

(a) Under 28 weeks (b) 32 to 35 weeks
(c) 37 to 39 weeks (d) 42 weeks and over

3. Describe the weights of the top 10% of the babies born with each gestation period. Explain your reasoning.

(a) 37 to 39 weeks (b) 42 weeks and over

4. For each gestation period, what is the probability that a baby will weigh between 6 and 9 pounds at birth?

(a) 32 to 35 weeks (b) 37 to 39 weeks
(c) 42 weeks and over

5. A birth weight of less than 3.3 pounds is classified by the NCHS as a "very low birth weight." What is the probability that a baby has a very low birth weight for each gestation period?

(a) Under 28 weeks (b) 32 to 35 weeks
(c) 37 to 39 weeks

5.4 Sampling Distributions and the Central Limit Theorem

What You Should Learn

- How to find sampling distributions and verify their properties
- How to interpret the Central Limit Theorem
- How to apply the Central Limit Theorem to find the probability of a sample mean

Sampling Distributions • The Central Limit Theorem • Probability and the Central Limit Theorem

Insight

Sample means can vary from one another and can also vary from the population mean. This type of variation is to be expected and is called *sampling error*.

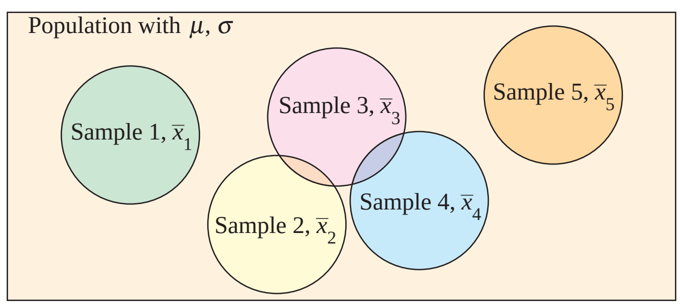
Sampling Distributions

In previous sections, you studied the relationship between the mean of a population and values of a random variable. In this section, you will study the relationship between a population mean and the means of samples taken from the population.

DEFINITION

A **sampling distribution** is the probability distribution of a sample statistic that is formed when samples of size n are repeatedly taken from a population. If the sample statistic is the sample mean, then the distribution is the **sampling distribution of sample means.**

For instance, consider the following Venn diagram. The rectangle represents a large population, and each circle represents a sample of size n. Because the sample entries can differ, the sample means can also differ. The mean of Sample 1 is $\bar{x}_1$; the mean of Sample 2 is $\bar{x}_2$; and so on. The sampling distribution of the sample means for samples of size n for this population consists of $\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$, and so on. If the samples are drawn with replacement, an infinite number of samples can be drawn from the population.



Note to Instructor

A good exercise that can be used in conjunction with the Venn diagram is to have each student randomly select a place in the random number table and write down the next five digits horizontally. Students can verify that the population of digits $\{0, 1, 2, \ldots, 9\}$ is uniform and has a mean of 4.5 and standard deviation of 2.87. Have each student calculate the mean of his or her sample and write that result on the board. Students can easily see that the sample means vary but are not dispersed as much as the population (range 0 to 9) is. Construct a histogram of the sample means; find the mean of these means and the standard deviation of the means. (With a TI-83, this takes little time even if only one student does the calculations.) Because the population standard deviation is known for this simulation, the results will be approximately normal.

Properties of Sampling Distributions of Sample Means

1. The mean of the sample means $\mu_{\bar{x}}$ is equal to the population mean μ.

$$\mu_{\bar{x}} = \mu$$

2. The standard deviation of the sample means $\sigma_{\bar{x}}$ is equal to the population standard deviation σ divided by the square root of n.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

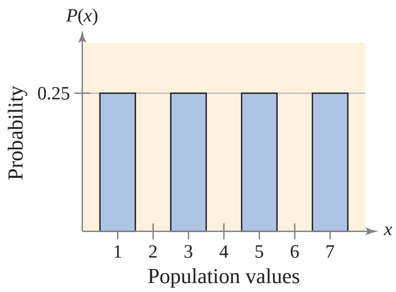
The standard deviation of the sampling distribution of the sample means is called the **standard error of the mean.**

EXAMPLE 1

A Sampling Distribution of Sample Means

You write the population values $\{1, 3, 5, 7\}$ on slips of paper and put them in a box. Then you randomly choose two slips of paper, with replacement. List all possible samples of size $n = 2$ and calculate the mean of each. These means form the sampling distribution of the sample means. Find the mean, variance, and standard deviation of the sample means. Compare your results with the mean $\mu = 4$, variance $\sigma^2 = 5$, and standard deviation $\sigma = \sqrt{5} \approx 2.236$ of the population.

Probability Histogram of Population of x



SOLUTION List all 16 samples of size 2 from the population and the mean of each sample.

Sample	Sample mean, $\bar{x}$	Sample	Sample mean, $\bar{x}$
1, 1	1	5, 1	3
1, 3	2	5, 3	4
1, 5	3	5, 5	5
1, 7	4	5, 7	6
3, 1	2	7, 1	4
3, 3	3	7, 3	5
3, 5	4	7, 5	6
3, 7	5	7, 7	7

Probability Distribution of Sample Means

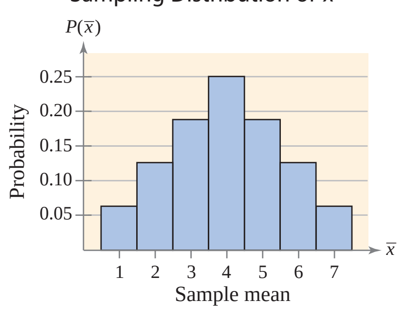
$\bar{x}$	f	Probability
1	1	0.0625
2	2	0.1250
3	3	0.1875
4	4	0.2500
5	3	0.1875
6	2	0.1250
7	1	0.0625

After constructing a probability distribution of the sample means, you can graph the sampling distribution using a probability histogram as shown at the left. Notice that the shape of the histogram is bell shaped and symmetric, similar to a normal curve. The mean, variance, and standard deviation of the 16 sample means are

$$\mu_{\bar{x}} = 4$$

$$(\sigma_{\bar{x}})^2 = \frac{5}{2} = 2.5 \quad \text{and} \quad \sigma_{\bar{x}} = \sqrt{\frac{5}{2}} = \sqrt{2.5} \approx 1.581.$$

Probability Histogram of Sampling Distribution of $\bar{x}$



These results satisfy the properties of sampling distributions because

$$\mu_{\bar{x}} = \mu = 4 \quad \text{and} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{5}}{\sqrt{2}} \approx \frac{2.236}{\sqrt{2}} \approx 1.581.$$

Try It Yourself 1

List all possible samples of $n = 3$, with replacement, from the population $\{1, 3, 5, 7\}$. Calculate the mean, variance, and standard deviation of the sample means. Compare these values with the corresponding population parameters.

a. *Form* all possible samples of size 3 and find the mean of each.
b. *Make* a probability distribution of the sample means and *find* the mean, variance, and standard deviation.
c. *Compare* the mean, variance, and standard deviation of the sample means with those for the population.

Answer: Page A36

Study Tip

Review Section 4.1 to find the mean and standard deviation of a probability distribution.

The Central Limit Theorem

The Central Limit Theorem forms the foundation for the inferential branch of statistics. This theorem describes the relationship between the sampling distribution of sample means and the population that the samples are taken from. The Central Limit Theorem is an important tool that provides the information you'll need to use sample statistics to make inferences about a population mean.

Note to Instructor

The sample mean $\bar{x}$ varies from sample to sample and is a random variable. As a random variable, it has a probability distribution, called the sampling distribution of the mean. Mention that other sample statistics, such as s^2, s, and $\hat{p}$, have different sampling distributions that will be studied in the next chapter.

The Central Limit Theorem

1. If samples of size n, where $n \geq 30$, are drawn from any population with a mean μ and a standard deviation σ, then the sampling distribution of sample means approximates a normal distribution. The greater the sample size, the better the approximation.
2. If the population itself is normally distributed, the sampling distribution of sample means is normally distributed for *any* sample size n.

In either case, the sampling distribution of sample means has a mean equal to the population mean.

$$\mu_{\bar{x}} = \mu \qquad \text{Mean}$$

The sampling distribution of sample means has a variance equal to $1/n$ times the variance of the population and a standard deviation equal to the population standard deviation divided by the square root of n.

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \qquad \text{Variance}$$

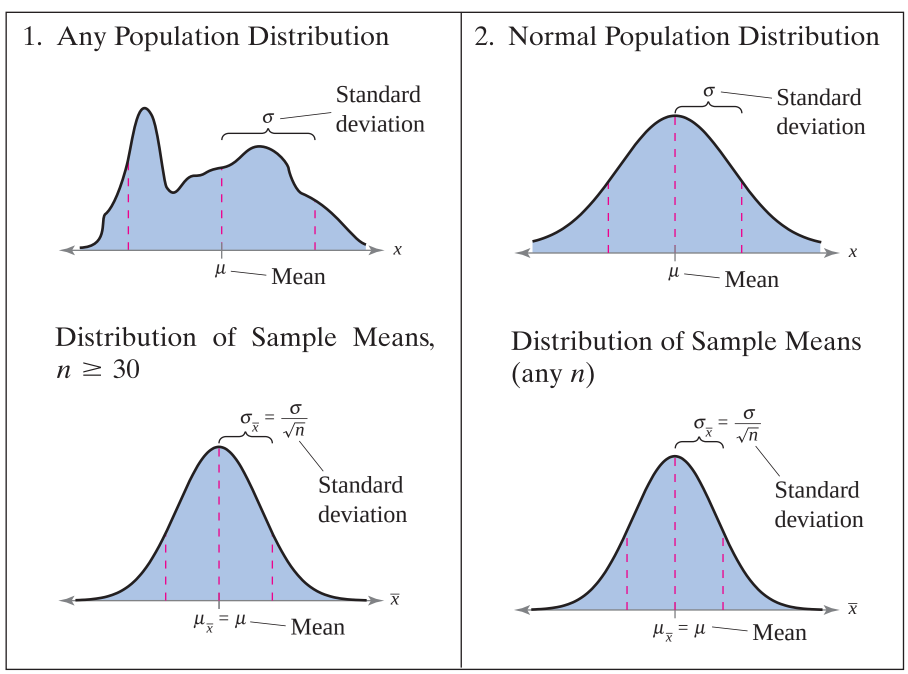
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad \text{Standard deviation}$$

The standard deviation of the sampling distribution of the sample means, $\sigma_{\bar{x}}$, is also called the **standard error of the mean.**

Insight

The distribution of sample means has the same mean as the population. But its standard deviation is less than the standard deviation of the population. This tells you that the distribution of sample means has the same center as the population, but it is not as spread out.

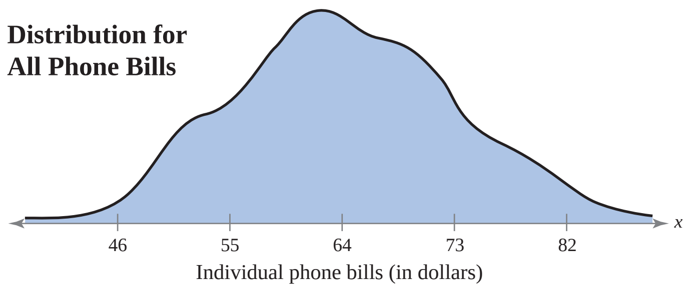
Moreover, the distribution of sample means becomes less and less spread out (tighter concentration about the mean) as the sample size n increases.



EXAMPLE 2

Interpreting the Central Limit Theorem

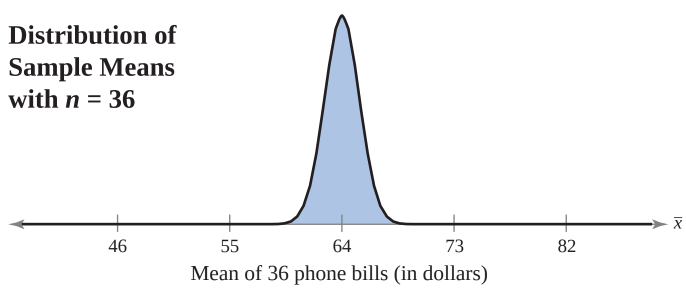
Phone bills for residents of a city have a mean of \$64 and a standard deviation of \$9, as shown in the following graph. Random samples of 36 phone bills are drawn from this population, and the mean of each sample is determined. Find the mean and standard error of the mean of the sampling distribution. Then sketch a graph of the sampling distribution of sample means.



SOLUTION The mean of the sampling distribution is equal to the population mean, and the standard error of the mean is equal to the population standard deviation divided by $\sqrt{n}$. So,

$$\mu_{\bar{x}} = \mu = 64 \quad \text{and} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{9}{\sqrt{36}} = 1.5.$$

Interpretation From the Central Limit Theorem, because the sample size is greater than 30, the sampling distribution can be approximated by a normal distribution with $\mu = \$64$ and $\sigma = \$1.50$, as shown in the graph below.



Try It Yourself 2

Suppose random samples of size 100 are drawn from the population in Example 2. Find the mean and standard error of the mean of the sampling distribution. Sketch a graph of the sampling distribution and compare it with the sampling distribution in Example 2.

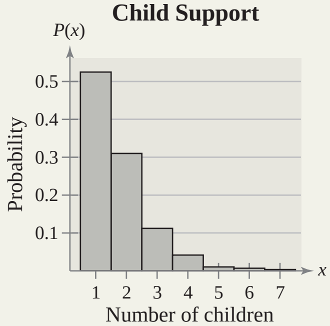
a. *Find* $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$.
b. *Identify* the sample size. If $n \geq 30$, *sketch* a normal curve with mean $\mu_{\bar{x}}$ and standard deviation $\sigma_{\bar{x}}$.
c. *Compare* the results with those in Example 2.

Answer: Page A36

Picturing the World

In a recent year, there were more than 5 million parents in the United States who received child support payments. The following histogram shows the distribution of children per custodial parent. The mean number of children was 1.7 and the standard deviation was 0.9. *(Adapted from U.S. Census Bureau)*

Child Support

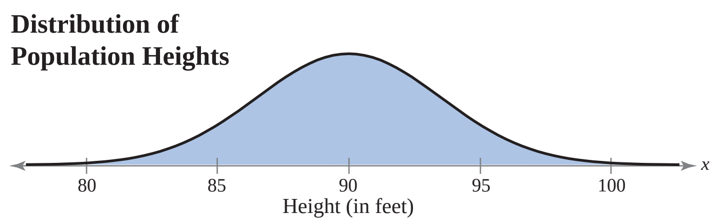


You randomly select 35 parents who receive child support and ask how many children in their custody are receiving child support payments. What is the probability that the mean of the sample is between 1.5 and 1.9 children?

EXAMPLE 3

Interpreting the Central Limit Theorem

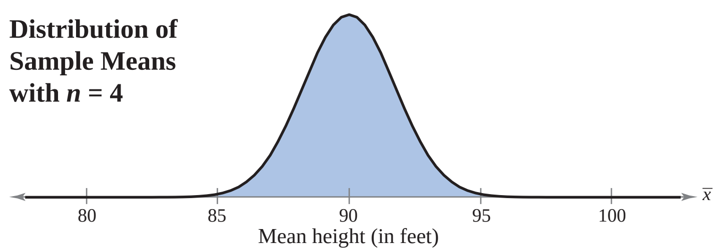
The heights of fully grown white oak trees are normally distributed, with a mean of 90 feet and standard deviation of 3.5 feet, as shown in the following graph. Random samples of size 4 are drawn from this population, and the mean of each sample is determined. Find the mean and standard error of the mean of the sampling distribution. Then sketch a graph of the sampling distribution of sample means.



SOLUTION The mean of the sampling distribution is equal to the population mean, and the standard error of the mean is equal to the population standard deviation divided by $\sqrt{n}$. So,

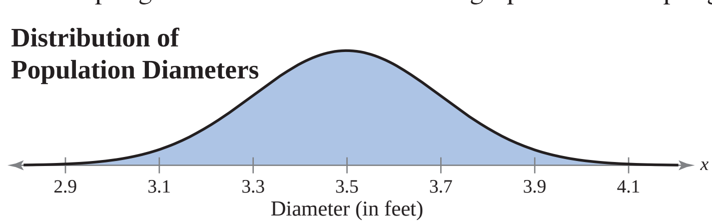
$$\mu_{\bar{x}} = \mu = 90 \text{ feet} \quad \text{and} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{3.5}{\sqrt{4}} = 1.75 \text{ feet.}$$

Interpretation From the Central Limit Theorem, because the population is normally distributed, the sampling distribution of the sample means is also normally distributed, as shown in the graph below.



Try It Yourself 3

The diameters of fully grown white oak trees are normally distributed, with a mean of 3.5 feet and a standard deviation of 0.2 foot, as shown in the graph below. Random samples of size 16 are drawn from this population, and the mean of each sample is determined. Find the mean and standard error of the mean of the sampling distribution. Then sketch a graph of the sampling distribution.



a. *Find* $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$.
b. *Sketch* a normal curve with mean $\mu_{\bar{x}}$ and standard deviation $\sigma_{\bar{x}}$.

Answer: Page A37

Probability and the Central Limit Theorem

Note to Instructor

For technology users, students need only calculate the standard error before using the normal CDF.

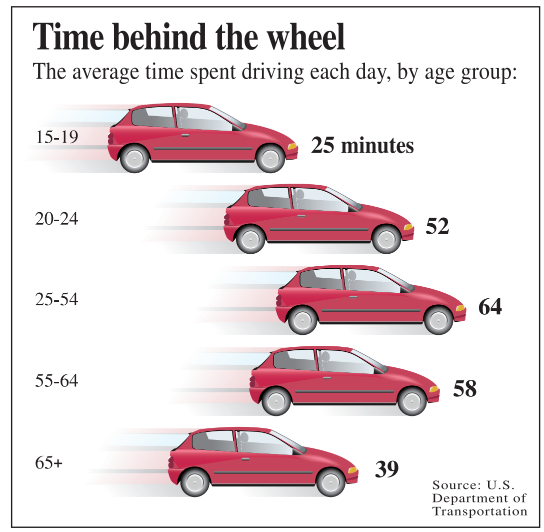
In Section 5.2, you learned how to find the probability that a random variable x will fall in a given interval of population values. In a similar manner, you can find the probability that a sample mean $\bar{x}$ will fall in a given interval of the $\bar{x}$ sampling distribution. To transform $\bar{x}$ to a z-score, you can use the formula

$$z = \frac{\text{Value} - \text{Mean}}{\text{Standard error}} = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}.$$

EXAMPLE 4

Finding Probabilities for Sampling Distributions

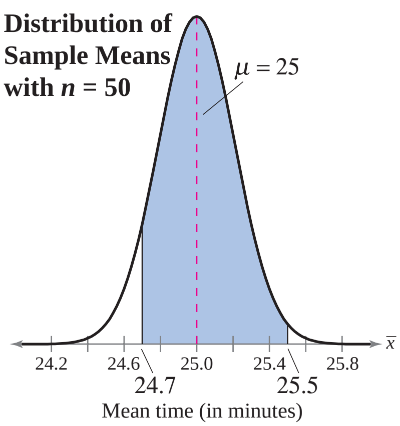
The graph at the right shows the length of time people spend driving each day. You randomly select 50 drivers ages 15 to 19. What is the probability that the mean time they spend driving each day is between 24.7 and 25.5 minutes? Assume that $\sigma = 1.5$ minutes.



Copyright 2003, USA TODAY. Reprinted with permission.

SOLUTION The sample size is greater than 30, so you can use the Central Limit Theorem to conclude that the distribution of sample means is approximately normal with a mean and a standard deviation of

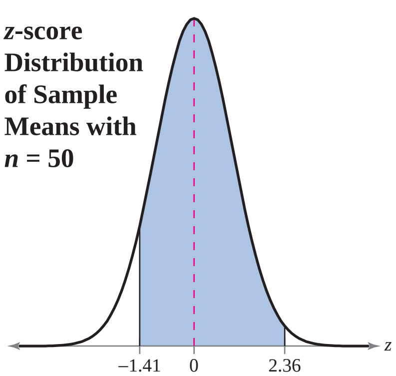
$$\mu_{\bar{x}} = \mu = 25 \text{ minutes} \quad \text{and} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.5}{\sqrt{50}} \approx 0.21213 \text{ minute}.$$



The graph of this distribution is shown at the left with a shaded area between 24.7 and 25.5 minutes. The z-scores that correspond to sample means of 24.7 and 25.5 minutes are

$$z_1 = \frac{24.7 - 25}{1.5/\sqrt{50}} \approx \frac{-0.3}{0.21213} \approx -1.41 \quad \text{and}$$

$$z_2 = \frac{25.5 - 25}{1.5/\sqrt{50}} \approx \frac{0.5}{0.21213} \approx 2.36.$$



So, the probability that the mean time the 50 people spend driving each day is between 24.7 and 25.5 minutes is

$$\begin{aligned} P(24.7 < \bar{x} < 25.5) &= P(-1.41 < z < 2.36) \\ &= P(z < 2.36) - P(z < -1.41) \\ &= 0.9909 - 0.0793 = 0.9116. \end{aligned}$$

Interpretation Of the samples of 50 drivers ages 15 to 19, 91.16% will have a mean driving time that is between 24.7 and 25.5 minutes, as shown in the graph at the left. This implies that, assuming the value of $\mu = 25$ is correct, only 8.84% of such sample means will lie outside the given interval.

Study Tip

Before you find probabilities for intervals of the sample mean $\overline{x}$, use the Central Limit Theorem to determine the mean and the standard deviation of the sampling distribution of the sample means. That is, calculate $\mu_{\overline{x}}$ and $\sigma_{\overline{x}}$.

Try It Yourself 4

You randomly select 100 drivers ages 15 to 19. What is the probability that the mean time they spend driving each day is between 24.7 and 25.5 minutes? Use $\mu = 25$ and $\sigma = 1.5$ minutes.

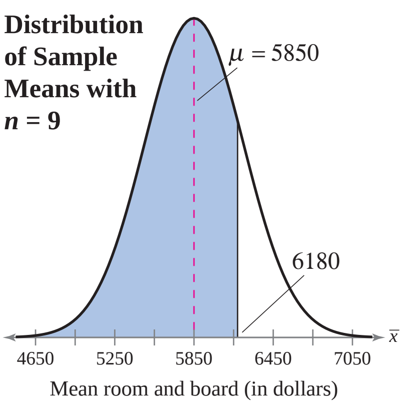
a. Use the Central Limit Theorem to *find* $\mu_{\overline{x}}$ and $\sigma_{\overline{x}}$ and *sketch* the sampling distribution of the sample means.
b. *Find the z-scores* that correspond to $\overline{x} = 24.7$ minutes and $\overline{x} = 25.5$ minutes.
c. *Find the cumulative area* that corresponds to each z-score and calculate the probability.

Answer: Page A37

EXAMPLE 5

Finding Probabilities for Sampling Distributions

The mean room and board expense per year at a four-year college is \$5850. You randomly select 9 four-year colleges. What is the probability that the mean room and board is less than \$6180? Assume that the room and board expenses are normally distributed, with a standard deviation of \$1125. *(Source: National Center for Education Statistics)*



SOLUTION Because the population is normally distributed, you can use the Central Limit Theorem to conclude that the distribution of sample means is normally distributed, with a mean of \$5850 and a standard deviation of \$375.

$$\mu_{\overline{x}} = \mu = 5850 \quad \text{and} \quad \sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1125}{\sqrt{9}} = 375$$

The graph of this distribution is shown at the left. The area to the left of \$6180 is shaded. The z-score that corresponds to \$6180 is

$$z = \frac{6180 - 5850}{1125/\sqrt{9}} = \frac{5600}{375} = 0.88.$$

So, the probability that the mean room and board expense is less than \$6180 is

$$P(\overline{x} < 6180) = P(z < 0.88) = 0.8106.$$

Interpretation So, 81.06% of such samples with $n = 9$ will have a mean less than \$6180 and 18.94% of these sample means will lie outside this interval.

Try It Yourself 5

The average sales price of a single-family house in the United States is \$243,756. You randomly select 12 single-family houses. What is the probability that the mean sales price is more than \$200,000? Assume that the sales prices are normally distributed with a standard deviation of \$44,000. *(Source: Federal Housing Finance Board)*

a. Use the Central Limit Theorem to *find* $\mu_{\overline{x}}$ and $\sigma_{\overline{x}}$ and *sketch* the sampling distribution of the sample means.
b. *Find the z-score* that corresponds to $\overline{x} = \$200{,}000$.
c. *Find the cumulative area* that corresponds to the z-score and calculate the probability.

Answer: Page A37

The Central Limit Theorem can also be used to investigate rare occurrences. A rare occurrence is one that occurs with a probability of less than 5%.

EXAMPLE 6

Finding Probabilities for x and $\bar{x}$

A bank auditor claims that credit card balances are normally distributed, with a mean of \$2870 and a standard deviation of \$900.

1. What is the probability that a randomly selected credit card holder has a credit card balance less than \$2500?
2. You randomly select 25 credit card holders. What is the probability that their mean credit card balance is less than \$2500?
3. Compare the probabilities from (1) and (2) and interpret your answer in terms of the auditor's claim.

SOLUTION

1. In this case, you are asked to find the probability associated with a certain value of the random variable x. The z-score that corresponds to $x = \$2500$ is

$$z = \frac{x - \mu}{\sigma} = \frac{2500 - 2870}{900} \approx -0.41.$$

So, the probability that the card holder has a balance less than \$2500 is

$$P(x < 2500) = P(z < -0.41) = 0.3409.$$

2. Here, you are asked to find the probability associated with a sample mean $\bar{x}$. The z-score that corresponds to $\bar{x} = \$2500$ is

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$
$$= \frac{2500 - 2870}{900/\sqrt{25}} = \frac{-370}{180} \approx -2.06.$$

So, the probability that the mean credit card balance of the 25 card holders is less than \$2500 is

$$P(\bar{x} < 2500) = P(z < -2.06) = 0.0197.$$

3. ***Interpretation*** Although there is a 34% chance that an individual will have a balance less than \$2500, there is only a 2% chance that the mean of a sample of 25 will have a balance less than \$2500. Because there is only a 2% chance that the mean of a sample of 25 will have a balance less than \$2500, this is a rare occurrence. So, it is possible that the sample is unusual, or it is possible that the auditor's claim that the mean is \$2870 is incorrect.

Study Tip

To find probabilities for individual members of a population with a normally distributed random variable x, use the formula

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

To find probabilities for the mean $\bar{x}$ of a sample size n, use the formula

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}.$$

Try It Yourself 6

A consumer price analyst claims that prices for sound-system receivers are normally distributed, with a mean of \$625 and a standard deviation of \$150. (1) What is the probability that a randomly selected receiver costs less than \$700? (2) You randomly select 10 receivers. What is the probability that their mean cost is less than \$700? (3) Compare these two probabilities.

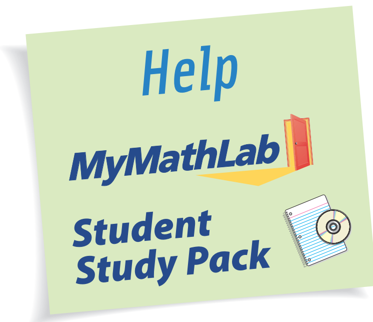
a. *Find* the z-scores that correspond to x and $\bar{x}$.
b. Use the Standard Normal Table to *find the probability* associated with each z-score.
c. *Compare* the probabilities and *interpret* your answer.

Answer: Page A37

Note to Instructor

You may want to tell students that the second formula can also be used to calculate z-scores for individual values. Consider a sample of $n = 1$ for an individual value.

5.4 Exercises



Building Basic Skills and Vocabulary

In Exercises 1–4, a population has a mean $\mu = 100$ and a standard deviation $\sigma = 15$. Find the mean and standard deviation of a sampling distribution of sample means with the given sample size n.

1. $n = 50$ **2.** $n = 100$

3. $n = 250$ **4.** $n = 1000$

True or False? In Exercises 5–8, determine whether the statement is true or false. If it is false, rewrite it so that it is a true statement.

5. As the size of a sample increases, the mean of the distribution of sample means increases.

6. As the size of a sample increases, the standard deviation of the distribution of sample means increases.

7. The shape of a sampling distribution is normal only if the shape of the population is normal.

8. If the size of a sample is at least 30, you can use z-scores to determine the probability that a sample mean falls in a given interval of the sampling distribution.

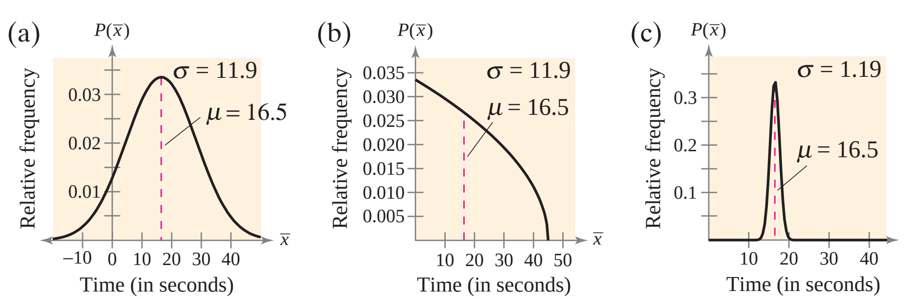
Verifying Properties of Sampling Distributions In Exercises 9 and 10, find the mean and standard deviation of the population. List all samples (with replacement) of the given size from that population. Find the mean and standard deviation of the sampling distribution and compare them with the mean and standard deviation of the population.

9. The number of movies that all four people in a family have seen in the past month is 4, 2, 8, and 0. Use a sample size of 3.

10. Four people in a carpool paid the following amounts for textbooks this semester: \$120, \$140, \$180, and \$220. Use a sample size of 2.

Graphical Analysis In Exercises 11 and 12, the graph of a population distribution is shown at the left with its mean and standard deviation. Assume that a sample size of 100 is drawn from each population. Decide which of the graphs labeled (a)–(c) would most closely resemble the sampling distribution of the sample means for each graph. Explain your reasoning.

11. The waiting time (in seconds) at a traffic signal during a red light



1. 100, 2.12 **2.** 100, 1.5

3. 100, 0.949 **4.** 100, 0.474

5. False. As the size of a sample increases, the mean of the distribution of sample means does not change.

6. False. As the size of the sample increases, the standard deviation of the distribution of sample means decreases.

7. False. The shape of a sampling distribution is normal if either $n \geq 30$ or the shape of the population is normal.

8. True

9. See Odd Answers, page A##.

10. {120 120, 120 140, 120 180, 120 220, 140 120, 140 140, 140 180, 140 220, 180 120, 180 140, 180 180, 180 220, 220 120, 220 140, 220 180, 220 220}

$\mu_{\bar{x}} = 165, \sigma_{\bar{x}} \approx 27.157$

$\mu = 165, \sigma = 38.406$

11. (c), because $\mu = 16.5$, $\sigma = 1.19$, and the graph approximates a normal curve.

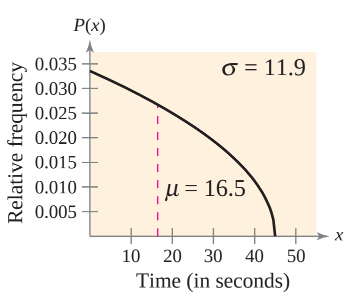


Figure for Exercise 11

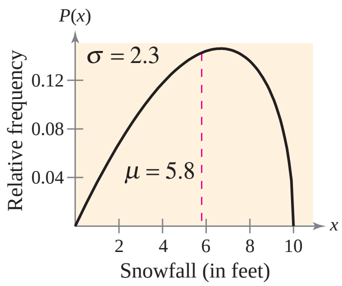
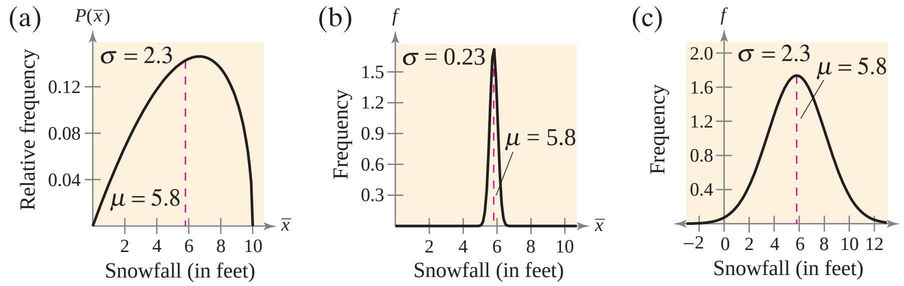


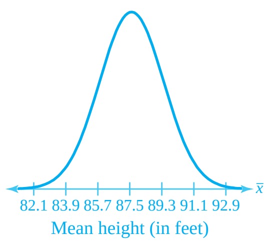
Figure for Exercise 12

12. The annual snowfall (in feet) for a central New York State county



12. See Selected Answers, page A##.

13. 87.5, 1.804

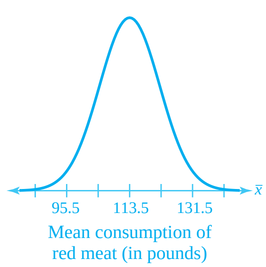


14. See Selected Answers, page A##.

15. 349, 1.26

16. See Selected Answers, page A##.

17. 113.5, 8.61



18. See Selected Answers, page A##.

19. See Odd Answers, page A##.

20. See Selected Answers, page A##.

Using and Interpreting Concepts

Using the Central Limit Theorem In Exercises 13–18, use the Central Limit Theorem to find the mean and standard error of the mean of the indicated sampling distribution. Then sketch a graph of the sampling distribution.

13. ***Heights of Trees*** The heights of fully grown sugar maple trees are normally distributed, with a mean of 87.5 feet and a standard deviation of 6.25 feet. Random samples of size 12 are drawn from the population and the mean of each sample is determined.

14. ***Fly Eggs*** The number of eggs a female house fly lays during her lifetime is normally distributed, with a mean of 800 eggs and a standard deviation of 100 eggs. Random samples of size 15 are drawn from this population and the mean of each sample is determined.

15. ***Digital Cameras*** The prices of digital cameras are normally distributed, with a mean of \$349 and a standard deviation of \$8. Random samples of size 40 are drawn from this population and the mean of each sample is determined.

16. ***Employees' Ages*** The ages of employees at a large corporation are normally distributed, with a mean of 47.2 years and a standard deviation of 3.6 years. Random samples of size 36 are drawn from this population and the mean of each sample is determined.

17. ***Red Meat Consumed*** The per capita consumption of red meat by people in the United States in a recent year was normally distributed, with a mean of 113.5 pounds and a standard deviation of 38.5 pounds. Random samples of size 20 are drawn from this population and the mean of each sample is determined. *(Adapted from U.S. Department of Agriculture)*

18. ***Soft Drinks*** The per capita consumption of soft drinks by people in the United States in a recent year was normally distributed, with a mean of 49.3 gallons and a standard deviation of 17.1 gallons. Random samples of size 25 are drawn from this population and the mean of each sample is determined. *(Adapted from U.S. Department of Agriculture)*

19. Repeat Exercise 13 for samples of size 24 and 36. What happens to the mean and standard deviation of the distribution of sample means as the size of the sample increases?

20. Repeat Exercise 14 for samples of size 30 and 45. What happens to the mean and to the standard deviation of the distribution of sample means as the size of the sample increases?

21. 0.0019
22. ≈ 0
23. 0.6319
24. 0.2349
25. ≈ 0
26. 0.0162
27. It is more likely to select a sample of 20 women with a mean height less than 70 inches because the sample of 20 has a higher probability.
28. It is more likely to select one man with a height less than 65 inches because the probability is greater.
29. Yes, it is very unlikely that you would have randomly sampled 40 cans with a mean equal to 127.9 ounces.
30. Yes, it is very unlikely that you would have randomly sampled 40 containers with a mean equal to 64.05 ounces.

Finding Probabilities In Exercises 21–26, find the probabilities.

21. ***Plumber Salaries*** The population mean annual salary for plumbers is $\mu = \$40{,}500$. A random sample of 42 plumbers is drawn from this population. What is the probability that the mean salary of the sample, $\bar{x}$, is less than \$38,000? Assume $\sigma = \$5600$. *(Adapted from Salary.com)*

22. ***Nurse Salaries*** The population mean annual salary for registered nurses is $\mu = \$45{,}500$. A sample of 35 registered nurses is randomly selected. What is the probability that the mean annual salary of the sample, $\bar{x}$, is less than \$42,000? Assume $\sigma = \$1700$. *(Adapted from Allied Physicians, Inc.)*

23. ***Gas Prices: New England*** During a certain week the mean price of gasoline in the New England region was $\mu = \$1.689$ per gallon. What is the probability that the mean price $\bar{x}$ for a sample of 32 randomly selected gas stations in that area was between \$1.684 and \$1.699 that week? Assume $\sigma = \$0.045$. *(Adapted from Energy Information Administration)*

24. ***Gas Prices: California*** During a certain week the mean price of gasoline in California was $\mu = \$2.029$ per gallon. A random sample of 38 gas stations is drawn from this population. What is the probability that $\bar{x}$, the mean price for the sample, was between \$2.034 and \$2.044? Assume $\sigma = \$0.049$. *(Adapted from Energy Information Administration)*

25. ***Heights of Women*** The mean height of women in the United States (ages 20–29) is $\mu = 64$ inches. A random sample of 60 women in this age group is selected. What is the probability that $\bar{x}$, the mean height for the sample, is greater than 66 inches? Assume $\sigma = 2.75$ inches. *(Source: National Center for Health Statistics)*

26. ***Heights of Men*** The mean height of men in the United States (ages 20–29) is $\mu = 69.2$ inches. A random sample of 60 men in this age group is selected. What is the probability that $\bar{x}$, the mean height for the sample, is greater than 70 inches? Assume $\sigma = 2.9$ inches. *(Source: National Center for Health Statistics)*

27. ***Which Is More Likely?*** Assume that the heights given in Exercise 25 are normally distributed. Are you more likely to randomly select one woman with a height less than 70 inches or are you more likely to select a sample of 20 women with a mean height less than 70 inches? Explain.

28. ***Which Is More Likely?*** Assume that the heights given in Exercise 26 are normally distributed. Are you more likely to randomly select one man with a height less than 65 inches or are you more likely to select a sample of 15 men with a mean height less than 65 inches? Explain.

29. ***Make a Decision*** A machine used to fill gallon-sized paint cans is regulated so that the amount of paint dispensed has a mean of 128 ounces and a standard deviation of 0.20 ounce. You randomly select 40 cans and carefully measure the contents. The sample mean of the cans is 127.9 ounces. Does the machine need to be reset? Explain your reasoning.

30. ***Make a Decision*** A machine used to fill pint-sized milk containers is regulated so that the amount of milk dispensed has a mean of 64 ounces and a standard deviation of 0.11 ounce. You randomly select 40 containers and carefully measure the contents. The sample mean of the containers is 64.05 ounces. Does the machine need to be reset? Explain your reasoning.

31. (a) 0.0008
(b) Claim is inaccurate.
(c) No, assuming the manufacturer's claim is true, because 96.25 is within 1 standard deviation of the mean for an individual board.

32. (a) 0.0179
(b) Claim is inaccurate.
(c) No, assuming the manufacturer's claim is true, because 10.21 is within 1 standard deviation of the mean for an individual carton.

33. (a) 0.0002
(b) Claim is inaccurate.
(c) No, assuming the manufacturer's claim is true, because 49,721 is within 1 standard deviation of the mean for an individual tire.

34. (a) 0.0668
(b) Claim is inaccurate.
(c) No, assuming the manufacturer's claim is true, because 37,650 is within 1 standard deviation of the mean for an individual brake pad.

31. ***Lumber Cutter*** Your lumber company has bought a machine that automatically cuts lumber. The seller of the machine claims that the machine cuts lumber to a mean length of 8 feet (96 inches) with a standard deviation of 0.5 inch. Assume the lengths are normally distributed. You randomly select 40 boards and find that the mean length is 96.25 inches.

(a) Assuming the seller's claim is correct, what is the probability the mean of the sample is 96.25 inches or more?
(b) Using your answer from part (a), what do you think of the seller's claim?
(c) Would it be unusual to have an individual board with a length of 96.25 inches? Why or why not?

32. ***Ice Cream Carton Weights*** A manufacturer claims that the mean weight of its ice cream cartons is 10 ounces with a standard deviation of 0.5 ounce. Assume the weights are normally distributed. You test 25 cartons and find their mean weight is 10.21 ounces.

(a) Assuming the manufacturer's claim is correct, what is the probability the mean of the sample is 10.21 ounces or more?
(b) Using your answer from part (a), what do you think of the manufacturer's claim?
(c) Would it be unusual to have an individual carton with a weight of 10.21 ounces? Why or why not?

33. ***Life of Tires*** A manufacturer claims that the life span of its tires is 50,000 miles. You work for a consumer protection agency and you are testing this manufacturer's tires. Assume the life spans of the tires are normally distributed. You select 100 tires at random and test them. The mean life span is 49,721 miles. Assume $\sigma = 800$ miles.

(a) Assuming the manufacturer's claim is correct, what is the probability the mean of the sample is 49,721 miles or less?
(b) Using your answer from part (a), what do you think of the manufacturer's claim?
(c) Would it be unusual to have an individual tire with a life span of 49,721 miles? Why or why not?

34. ***Brake Pads*** A brake pad manufacturer claims its brake pads will last for 38,000 miles. You work for a consumer protection agency and you are testing this manufacturer's brake pads. Assume the life spans of the brake pads are normally distributed. You randomly select 50 brake pads. In your tests, the mean life of the brake pads is 37,650 miles. Assume $\sigma = 1000$ miles.

(a) Assuming the manufacturer's claim is correct, what is the probability the mean of the sample is 37,650 miles or less?
(b) Using your answer from part (a), what do you think of the manufacturer's claim?
(c) Would it be unusual to have an individual brake pad last for 37,650 miles? Why or why not?

35. Yes, because of the relatively large z-score (2.12).

36. It is very unlikely the machine is calibrated to produce a bolt with a mean of 4 inches.

37. ≈ 0

38. ≈ 1

Extending Concepts

35. *SAT Scores* The average math SAT score is 500 with a standard deviation of 100. A particular high school claims that its students have unusually high math SAT scores. A random sample of 50 students from this school was selected, and the mean math SAT score was 530. Is the high school justified in its claim? Explain.

36. *Machine Calibrations* A machine in a manufacturing plant is calibrated to produce a bolt that has a mean diameter of 4 inches and a standard deviation of 0.5 inch. An engineer takes a random sample of 100 bolts from this machine and finds the mean diameter is 4.2 inches. What are some possible consequences from these findings?

Finite Correction Factor The formula for the standard error of the mean

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

given in the Central Limit Theorem is based on an assumption that the population has infinitely many members. This is the case whenever sampling is done with replacement (each member is put back after it is selected) because the sampling process could be continued indefinitely. The formula is also valid if the sample size is small in comparison to the population. However, when sampling is done without replacement and the sample size n is more than 5% of the finite population of size N, there is a finite number of possible samples. A **finite correction factor,**

$$\sqrt{\frac{N-n}{N-1}}$$

should be used to adjust the standard error. The sampling distribution of the sample means will be normal with a mean equal to the population mean, and the standard error of the mean will be

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}.$$

In Exercises 37 and 38, determine if the finite correction factor should be used. If so, use it in your calculations when you find the probability.

37. *Gas Prices* In a sample of 800 gas stations, the mean price for regular gasoline at the pump was \$1.688 per gallon and the standard deviation was \$0.009 per gallon. A random sample of size 55 is drawn from this population. What is the probability that the mean price per gallon is less than \$1.683? *(Adapted from U.S. Department of Energy)*

38. *Old Faithful* In a sample of 500 eruptions of the Old Faithful geyser at Yellowstone National Park, the mean duration of the eruptions was 3.32 minutes and the standard deviation was 1.09 minutes. A random sample of size 30 is drawn from this population. What is the probability that the mean duration of eruptions is between 2.5 minutes and 4 minutes? *(Adapted from Yellowstone National Park)*

5.5 Normal Approximations to Binomial Distributions

What You Should Learn

- How to decide when the normal distribution can approximate the binomial distribution
- How to find the correction for continuity
- How to use the normal distribution to approximate binomial probabilities

Approximating a Binomial Distribution • Correction for Continuity • Approximating Binomial Probabilities

Approximating a Binomial Distribution

In Section 4.2, you learned how to find binomial probabilities. For instance, if a surgical procedure has an 85% chance of success and a doctor performs the procedure on 10 patients, it is easy to find the probability of exactly two successful surgeries.

But what if the doctor performs the surgical procedure on 150 patients and you want to find the probability of *fewer than 100* successful surgeries? To do this using the techniques described in Section 4.2, you would have to use the binomial formula 100 times and find the sum of the resulting probabilities. This approach is not practical, of course. A better approach is to use a normal distribution to approximate the binomial distribution.

Normal Approximation to a Binomial Distribution

If $np \geq 5$ and $nq \geq 5$, then the binomial random variable x is approximately normally distributed, with mean

$$\mu = np$$

and standard deviation

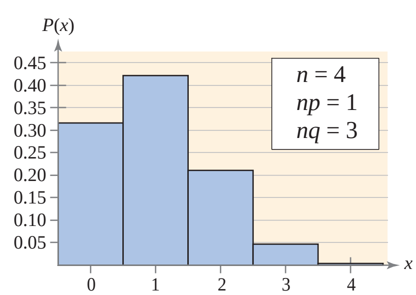
$$\sigma = \sqrt{npq}.$$

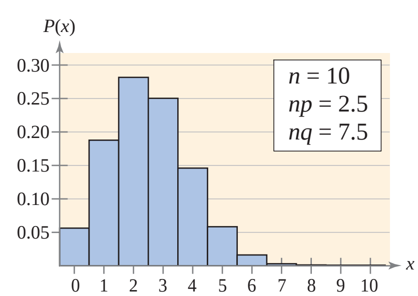
Study Tip

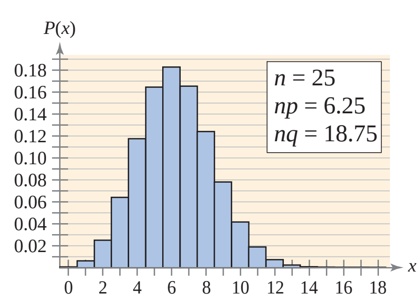
Properties of a binomial experiment

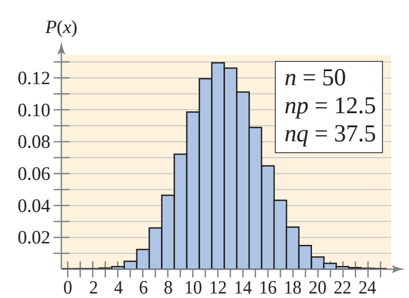
- n independent trials
- Two possible outcomes: success or failure
- Probability of success is p; probability of failure is $1 - p = q$
- p is constant for each trial

To see why this result is valid, look at the following binomial distributions for $p = 0.25$ and $n = 4$, $n = 10$, $n = 25$, and $n = 50$. Notice that as n increases, the histogram approaches a normal curve.









EXAMPLE 1

Approximating the Binomial Distribution

Two binomial experiments are listed. Decide whether you can use the normal distribution to approximate x, the number of people who reply yes. If you can, find the mean and standard deviation. If you cannot, explain why. *(Source: Marist College Institute for Public Opinion)*

1. Thirty-four percent of people in the United States say that they are likely to make a New Year's resolution. You randomly select 15 people in the United States and ask each if he or she is likely to make a New Year's resolution.
2. Six percent of people in the United States who made a New Year's resolution resolved to exercise more. You randomly select 65 people in the United States who made a resolution and ask each if he or she resolved to exercise more.

SOLUTION

1. In this binomial experiment, $n = 15$, $p = 0.34$, and $q = 0.66$. So,

$$np = (15)(0.34) = 5.1$$

and

$$nq = (15)(0.66) = 9.9.$$

Because np and nq are greater than 5, you can use the normal distribution with

$$\mu = 5.10$$

and

$$\sigma = \sqrt{npq} = \sqrt{15 \cdot 0.34 \cdot 0.66} \approx 1.83$$

to approximate the distribution of x.

2. In this binomial experiment, $n = 65$, $p = 0.06$, and $q = 0.94$. So,

$$np = (65)(0.06) = 3.9$$

and

$$nq = (65)(0.94) = 61.1.$$

Because $np < 5$, you cannot use the normal distribution to approximate the distribution of x.

Try It Yourself 1

Consider the following binomial experiment. Decide whether you can use the normal distribution to approximate x, the number of people who reply yes. If you can, find the mean and standard deviation. If you cannot, explain why. *(Source: Marist College Institute for Public Opinion)*

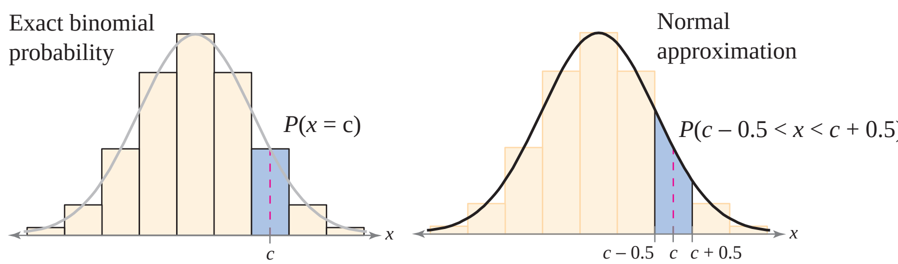
Sixty-one percent of people in the United States who made a New Year's resolution last year kept it. You randomly select 70 people in the United States who made a resolution last year and ask each if he or she kept the resolution.

a. *Identify* n, p, and q.
b. *Find* the products np and nq.
c. *Decide* whether you can use the normal distribution to approximate x.
d. *Find* the mean μ and standard deviation σ, if appropriate. *Answer: Page A37*

Correction for Continuity

The binomial distribution is discrete and can be represented by a probability histogram. To calculate *exact* binomial probabilities, you can use the binomial formula for each value of x and add the results. Geometrically, this corresponds to adding the areas of bars in the probability histogram. Remember that each bar has a width of one unit and x is the midpoint of the interval.

When you use a *continuous* normal distribution to approximate a binomial probability, you need to move 0.5 unit to the left and right of the midpoint to include all possible x-values in the interval. When you do this, you are making a **correction for continuity.**



Note to Instructor

For technology users who are not limited to $n = 20$ in the table, many more binomial problems can be calculated without using a normal distribution approximation. However, students should be shown that even technology has limitations. The TI-83 cannot calculate the cumulative binomial probability for $n = 10{,}000$, $p = 0.4$, and $x = 9000$, but that probability can be calculated using a normal approximation. Likewise, depending on the version of MINITAB or Excel you are using, there are memory limitations for the binomial distribution.

Study Tip

To use a correction for continuity, simply subtract 0.5 from the lowest value and add 0.5 to the highest.

EXAMPLE 2

Using a Correction for Continuity

Use a correction for continuity to convert each of the following binomial intervals to a normal distribution interval.

1. The probability of getting between 270 and 310 successes, inclusive
2. The probability of at least 158 successes
3. The probability of getting less than 63 successes

SOLUTION

1. The discrete midpoint values are 270, 271, . . . , 310. The corresponding interval for the continuous normal distribution is

   $269.5 < x < 310.5.$

2. The discrete midpoint values are 158, 159, 160, ... The corresponding interval for the continuous normal distribution is

   $x > 157.5.$

3. The discrete midpoint values are . . . , 60, 61, 62. The corresponding interval for the continuous normal distribution is

   $x < 62.5.$

Try It Yourself 2

Use a correction for continuity to convert each of the following binomial intervals to a normal distribution interval.

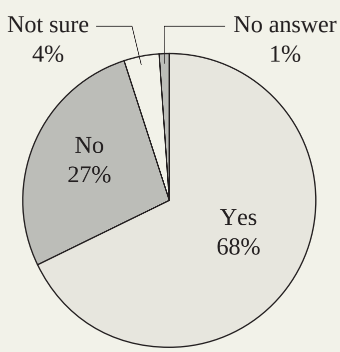
1. The probability of getting between 57 and 83 successes, inclusive
2. The probability of getting at most 54 successes

a. List the *midpoint values* for the binomial probability.
b. Use a *correction for continuity* to write the normal distribution interval.

Answer: Page A37

Picturing the World

In a survey of U.S. adults, people were asked if the law should allow doctors to aid dying patients who want to end their lives. The results of the survey are shown in the following pie chart. *(Source: The Harris Poll)*



Assume that this Harris Poll is a true indication of the proportion of the population who believe in assisted death for terminally ill patients.

If you sampled 50 adults at random, what is the probability that between 32 and 36, inclusive, would believe in assisted death?

Approximating Binomial Probabilities

GUIDELINES

Using the Normal Distribution to Approximate Binomial Probabilities

In Words	*In Symbols*
1. Verify that the binomial distribution applies.	Specify n, p, and q.
2. Determine if you can use the normal distribution to approximate x, the binomial variable.	Is $np \geq 5$? Is $nq \geq 5$?
3. Find the mean μ and standard deviation σ for the distribution.	$\mu = np$ $\sigma = \sqrt{npq}$
4. Apply the appropriate continuity correction. Shade the corresponding area under the normal curve.	Add or subtract 0.5 from endpoints.
5. Find the corresponding z-score(s).	$z = \dfrac{x - \mu}{\sigma}$
6. Find the probability.	Use the Standard Normal Table.

EXAMPLE 3

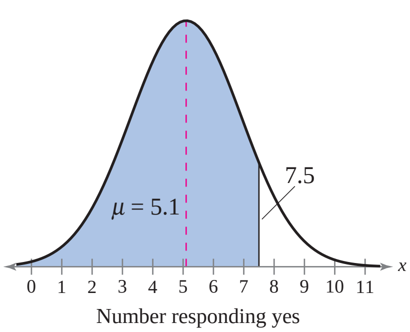
Approximating a Binomial Probability

Thirty-four percent of people in the United States say that they are likely to make a New Year's resolution. You randomly select 15 people in the United States and ask each if he or she is likely to make a New Year's resolution. What is the probability that fewer than eight of them respond yes? *(Source: Marist College Institute for Public Opinion)*

SOLUTION From Example 1, you know that you can use a normal distribution with $\mu = 5.1$ and $\sigma \approx 1.83$ to approximate the binomial distribution. Remember to apply the continuity correction for the value of x. In the binomial distribution, the possible midpoint values for "fewer than 8" are

$$\ldots 5, 6, 7.$$

To use the normal distribution, add 0.5 to the right-hand boundary 7 to get $x = 7.5$. The graph at the left shows a normal curve with $\mu = 5.1$ and $\sigma \approx 1.83$ and a shaded area to the left of 7.5. The z-score that corresponds to $x = 7.5$ is



$$z = \frac{7.5 - 5.1}{1.83}$$

$$\approx 1.31.$$

Using the Standard Normal Table,

$$P(z < 1.31) = 0.9049.$$

Interpretation The probability that fewer than eight people respond yes is approximately 0.9049, or about 91%.

Try It Yourself 3

Sixty-one percent of people in the United States who made a New Year's resolution last year kept it. You randomly select 70 people in the United States who made a resolution last year and ask each if he or she kept the resolution. What is the probability that more than 50 respond yes? (See Try It Yourself 1.) *(Source: Marist College Institute for Public Opinion)*

a. *Determine* whether you can use the normal distribution to approximate the binomial variable (see part c of Try It Yourself 1).
b. *Find* the mean μ and the standard deviation σ for the distribution (see part d of Try It Yourself 1).
c. *Apply* the appropriate continuity correction and sketch a graph.
d. *Find* the corresponding z-score.
e. *Use* the Standard Normal Table to find the area to the left of z and calculate the probability.

Answer: Page A37

EXAMPLE 4

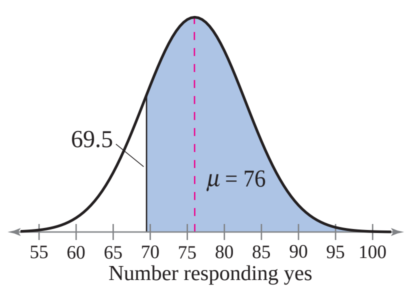
Approximating a Binomial Probability

Thirty-eight percent of people in the United States admit that they snoop in other people's medicine cabinets. You randomly select 200 people in the United States and ask each if he or she snoops in other people's medicine cabinets. What is the probability that at least 70 will say yes? *(Source: USA TODAY)*

SOLUTION Because $np = 200 \cdot 0.38 = 76$ and $nq = 200 \cdot 0.62 = 124$, the binomial variable x is approximately normally distributed with

$$\mu = np = 76 \quad \text{and} \quad \sigma = \sqrt{200 \cdot 0.38 \cdot 0.62} \approx 6.86.$$

Using the correction for continuity, you can rewrite the discrete probability $P(x \geq 70)$ as the continuous probability $P(x \geq 69.5)$. The graph shows a normal curve with $\mu = 76$ and $\sigma = 6.86$ and a shaded area to the right of 69.5. The z-score that corresponds to 69.5 is $z = (69.5 - 76)/6.86 \approx -0.95$. So, the probability that at least 70 will say yes is



$$\begin{aligned} P(x \geq 69.5) &= P(z \geq -0.95) \\ &= 1 - P(z \leq -0.95) \\ &= 1 - 0.1711 = 0.8289. \end{aligned}$$

Study Tip

In a discrete distribution, there is a difference between $P(x \geq c)$ and $P(x > c)$. This is true because the probability that x is exactly c is not zero. In a continuous distribution, however, there is no difference between $P(x \geq c)$ and $P(x > c)$ because the probability that x is exactly c is zero.

Try It Yourself 4

What is the probability that at most 85 people will say yes?

a. *Determine* whether you can use the normal distribution to approximate the binomial variable (see Example 4).
b. *Find* the mean μ and the standard deviation σ for the distribution.
c. *Apply* a continuity correction to rewrite $P(x \leq 85)$ and sketch a graph.
d. *Find* the corresponding z-score.
e. *Use* the Standard Normal Table to *find the area* to the left of z and calculate the probability.

Answer: Page A37

EXAMPLE 5

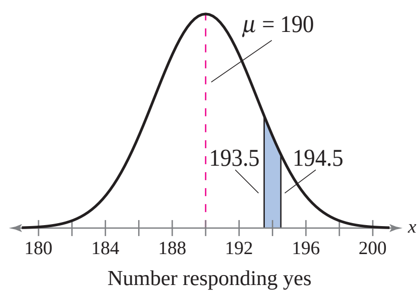
Approximating a Binomial Probability

A survey reports that 95% of Internet users use Microsoft Internet Explorer as their browser. You randomly select 200 Internet users and ask each whether he or she uses Microsoft Internet Explorer as his or her browser. What is the probability that exactly 194 will say yes? *(Source: OneStat.com)*

SOLUTION Because $np = 200 \cdot 0.95 = 190$ and $nq = 200 \cdot 0.05 = 10$, the binomial variable x is approximately normally distributed with

$$\mu = np = 190 \quad \text{and} \quad \sigma = \sqrt{200 \cdot 0.95 \cdot 0.05} \approx 3.08.$$

Using the correction for continuity, you can rewrite the discrete probability $P(x = 194)$ as the continuous probability $P(193.5 < x < 194.5)$. The following graph shows a normal curve with $\mu = 190$ and $\sigma = 3.08$ and a shaded area between 193.5 and 194.5.



The z-scores that correspond to 193.5 and 194.5 are

$$z_1 = \frac{193.5 - 190}{3.08} \approx 1.14 \quad \text{and} \quad z_2 = \frac{194.5 - 190}{3.08} \approx 1.46.$$

So, the probability that exactly 194 Internet users will say they use Microsoft Internet Explorer is

$$\begin{aligned} P(193.5 < x < 194.5) &= P(1.14 < z < 1.46) \\ &= P(z < 1.46) - P(z < 1.14) \\ &= 0.9279 - 0.8729 \\ &= 0.0550. \end{aligned}$$

Interpretation There is a probability of about 0.06 that exactly 194 of the Internet users will say they use Microsoft Internet Explorer.

Try It Yourself 5

What is the probability that exactly 191 people will say yes?

a. *Determine* whether you can use the normal distribution to approximate the binomial variable (see Example 5).
b. *Find* the mean μ and the standard deviation σ for the distribution.
c. *Apply* a continuity correction to rewrite $P(x = 191)$ and sketch a graph.
d. *Find* the corresponding z-scores.
e. *Use* the Standard Normal Table to *find the area* to the left of each z-score and calculate the probability.

Answer: Page A37

Note to Instructor

You may want to have students calculate the probability using the binomial formula from Chapter 4 and compare results.

$$P(x = 194) = {}_{200}C_{194}(0.95)^{194}(0.05)^{6}$$

The TI-83 gives 0.061400531.

5.5 Exercises

Building Basic Skills and Vocabulary

In Exercises 1–4, the sample size n, probability of success p, and probability of failure q are given for a binomial experiment. Decide whether you can use the normal distribution to approximate the random variable x.

1. $n = 20,\ p = 0.80,\ q = 0.20$

2. $n = 12,\ p = 0.60,\ q = 0.40$

3. $n = 15,\ p = 0.65,\ q = 0.35$

4. $n = 18,\ p = 0.85,\ q = 0.15$

Approximating a Binomial Distribution In Exercises 5–8, a binomial experiment is given. Decide whether you can use the normal distribution to approximate the binomial distribution. If you can, find the mean and standard deviation. If you cannot, explain why.

5. ***Credit Card Contract*** A survey of U.S. adults found that 44% read every word of a credit card contract. You ask 10 adults selected at random if he or she reads every word of a credit card contract. *(Source: USA TODAY)*

6. ***Organ Donors*** A survey of U.S. adults found that 63% would want their organs transplanted into a patient who needs them if they were killed in an accident. You randomly select 20 adults and ask each if he or she would want their organs transplanted into a patient who needs them if they were killed in an accident. *(Source: USA TODAY)*

7. ***Prostate Cancer*** In a recent year, the American Cancer Society said that the five-year survival rate for all men diagnosed with prostate cancer was 97%. You randomly select 10 men who were diagnosed with prostate cancer and calculate their five-year survival rate. *(Source: American Cancer Society)*

8. ***Work Weeks*** A survey of workers in the United States found that 8.6% work fewer than 40 hours per week. You randomly select 30 workers in the United States and ask each if he or she works fewer than 40 hours per week.

In Exercises 9–12, match the binomial probability with the correct statement.

Probability	Statement
9. $P(x \geq 45)$	(a) P(there are fewer than 45 successes)
10. $P(x \leq 45)$	(b) P(there are at most 45 successes)
11. $P(x < 45)$	(c) P(there are more than 45 successes)
12. $P(x > 45)$	(d) P(there are at least 45 successes)

In Exercises 13–16, use the correction for continuity and match the binomial probability statement with the corresponding normal distribution statement.

Binomial Probability	Normal Probability
13. $P(x > 89)$	(a) $P(x > 89.5)$
14. $P(x \geq 89)$	(b) $P(x < 88.5)$
15. $P(x \leq 89)$	(c) $P(x \leq 89.5)$
16. $P(x < 89)$	(d) $P(x \geq 88.5)$

Help

MyMathLab

Student Study Pack

1. Cannot use normal distribution.
2. Cannot use normal distribution.
3. Can use normal distribution.
4. Cannot use normal distribution.
5. Cannot use normal distribution because $np < 5$.
6. Can use normal distribution. $\mu = 12.6$, $\sigma = 2.159$
7. Cannot use normal distribution because $nq < 5$.
8. Cannot use normal distribution because $np < 5$.
9. d
10. b
11. a
12. c
13. a
14. d
15. c
16. b

17. Binomial: 0.549; Normal: 0.5463

18. Binomial: 0.19; Normal: 0.1875

19. Cannot use normal distribution because $np < 5$.

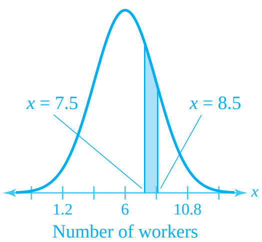
(a) 0.0000199 (b) 0.000023

(c) 0.999977 (d) 0.1635

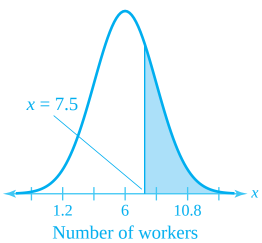
20. See Selected Answers, page A##.

21. Can use normal distribution.

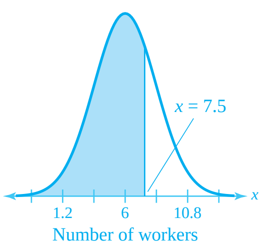
(a) 0.1174



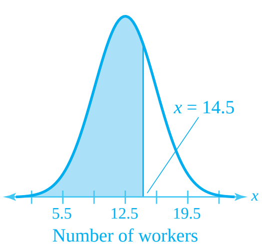
(b) 0.2643



(c) 0.7357



(d) 0.7190

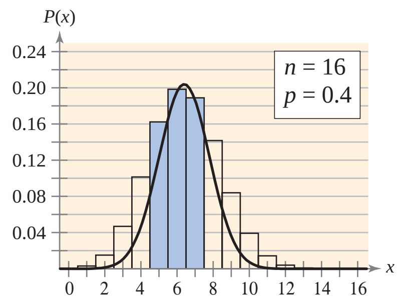


22. See Selected Answers, page A##.

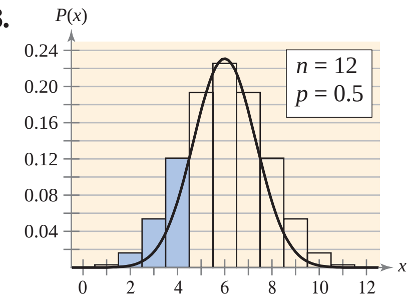
Using and Interpreting Concepts

Graphical Analysis In Exercises 17 and 18, write the binomial probability and the normal probability for the shaded region of the graph. Find the value of each probability and compare the results.

17.



18.



Approximating Binomial Probabilities In Exercises 19–24, decide whether you can use the normal distribution to approximate the binomial distribution. If you can, use the normal distribution to approximate the indicated probabilities and sketch their graphs. If you cannot, explain why and use the binomial distribution to find the indicated probabilities.

19. ***Blood Type O$^-$*** Seven percent of people in the United States have type O$^-$ blood. You randomly select 30 people in the United States and ask them if their blood type is O$^-$. *(Source: American Association of Blood Banks)*

(a) Find the probability that exactly 10 people say they have O$^-$ blood.

(b) Find the probability that at least 10 people say they have O$^-$ blood.

(c) Find the probability that fewer than 10 people say they have O$^-$ blood.

(d) A blood drive would like to get at least five donors with O$^-$ blood. There are 100 donors. What is the probability that there will not be enough O$^-$ blood donors?

20. ***Blood Type A$^+$*** Thirty-four percent of people in the United States have type A$^+$ blood. You randomly select 32 people in the United States and ask them if their blood type is A$^+$. *(Source: American Association of Blood Banks)*

(a) Find the probability that exactly 12 people say they have A$^+$ blood.

(b) Find the probability that at least 12 people say they have A$^+$ blood.

(c) Find the probability that fewer than 12 people say they have A$^+$ blood.

(d) A blood drive would like to get at least 60 donors with A$^+$ blood. There are 150 donors. What is the probability that there will not be enough A$^+$ blood donors?

21. ***Public Transportation*** Five percent of workers in the United States use public transportation to get to work. You randomly select 120 workers and ask them if they use public transportation to get to work. *(Source: U.S. Census Bureau)*

(a) Find the probability that exactly eight workers will say yes.

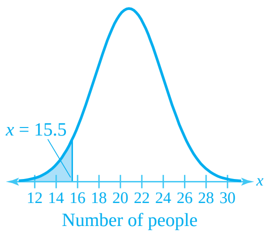
(b) Find the probability that at least eight workers will say yes.

(c) Find the probability that fewer than eight workers will say yes.

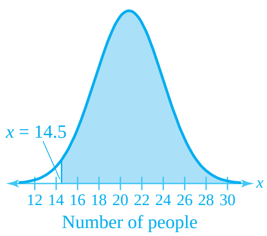
(d) A transit authority offers discount rates to companies that have at least 15 employees who use public transportation to get to work. There are 250 employees in a company. What is the probability that the company will not get the discount?

23. Can use normal distribution.

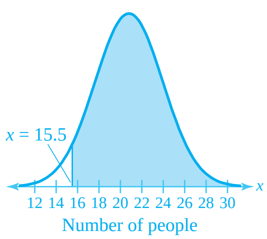
(a) 0.0465



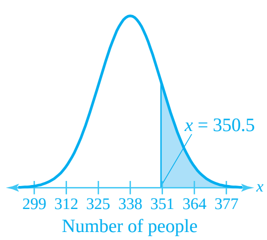
(b) 0.9767



(c) 0.9535



(d) 0.1635



24. Cannot use normal distribution because $np < 5$.

(a) 0.99987

(b) 0.00251

(c) 0.00013

(d) 0.230

25. (a) $np = 6 \geq 5$

$nq = 19 \geq 5$

(b) 0.121

(c) No, because the z-score is within one standard deviation of the mean.

22. *College Graduates* Thirty-one percent of workers in the United States are college graduates. You randomly select 50 workers and ask each if he or she is a college graduate. *(Source: U.S. Bureau of Labor Statistics)*

(a) Find the probability that exactly 14 workers are college graduates.

(b) Find the probability that at least 14 workers are college graduates.

(c) Find the probability that fewer than 14 workers are college graduates.

(d) A committee is looking for 30 working college graduates to volunteer at a career fair. The committee randomly selects 150 workers. What is the probability that there will not be enough college graduates?

23. *Favorite Cookie* Fifty-two percent of adults say chocolate chip is their favorite cookie. You randomly select 40 adults and ask each if chocolate chip is his or her favorite cookie. *(Source: WEAREVER)*

(a) Find the probability that at most 15 people say chocolate chip is their favorite cookie.

(b) Find the probability that at least 15 people say chocolate chip is their favorite cookie.

(c) Find the probability that more than 15 people say chocolate chip is their favorite cookie.

(d) A community bake sale has prepared 350 chocolate chip cookies. The bake sale attracts 650 customers, and they each buy one cookie. What is the probability there will not be enough chocolate chip cookies?

24. *Long Work Weeks* A survey of workers in the United States found that 2.9% work more than 70 hours per week. You randomly select 10 workers in the U.S. and ask each if he or she works more than 70 hours per week.

(a) Find the probability that at most three people say they work more than 70 hours per week.

(b) Find the probability that at least three people say they work more than 70 hours per week.

(c) Find the probability that more than three people say they work more than 70 hours per week.

(d) A large company is concerned about overworked employees who work more than 70 hours per week. The company randomly selects 50 employees. What is the probability there will be no employee working more than 70 hours?

25. *Bigger Home* A survey of homeowners in the United States found that 24% feel their home is too small for their family. You randomly select 25 homeowners and ask them if they feel their home is too small for their family.

(a) Verify that the normal distribution can be used to approximate the binomial distribution.

(b) Find the probability that more than eight homeowners say their home is too small for their family.

(c) Is it unusual for 8 out of 25 homeowners to say their home is too small? Why or why not?

26. (a) $np = 32 \geq 5; nq = 8 \geq 5$

(b) 0.0150

(c) Yes, because the z-score is more than two standard deviations from the mean.

27. Highly unlikely. Answers will vary.

28. Probable. Answers will vary.

29. 0.1020

30. 0.1736

26. ***Driving to Work*** A survey of workers in the United States found that 80% rely on their own vehicle to get to work. You randomly select 40 workers and ask them if they rely on their own vehicle to get to work.

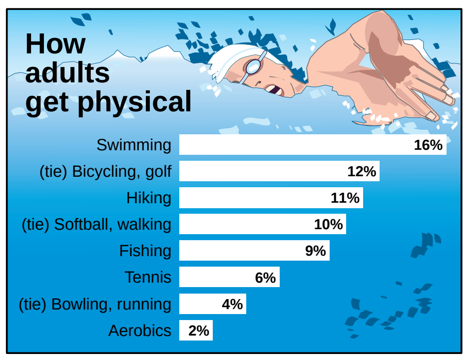
(a) Verify that the normal distribution can be used to approximate the binomial distribution.

(b) Find the probability that at most 26 workers say they rely on their own vehicle to get to work.

(c) Is it unusual for 26 out of 40 workers to say they rely on their own vehicle to get to work? Why or why not?

Extending Concepts

Getting Physical In Exercises 27 and 28, use the following information. The graph shows the results of a survey of adults in the United States ages 33 to 51 who were asked if they participated in a sport. Seventy percent of adults said they regularly participate in at least one sport, and they gave their favorite sport.



27. You randomly select 250 people in the United States ages 33 to 51 and ask each if he or she regularly participates in at least one sport. You find that 60% say no. How likely is this result? Do you think the sample is a good one? Explain your reasoning.

28. You randomly select 300 people in the United States ages 33 to 51 and ask each if he or she regularly participates in at least one sport. Of the 200 who say yes, 9% say they participate in hiking. How likely is this result? Is the sample a good one? Explain your reasoning.

Testing a Drug In Exercises 29 and 30, use the following information. A drug manufacturer claims that a drug cures a rare skin disease 75% of the time. The claim is checked by testing the drug on 100 patients. If at least 70 patients are cured, the claim will be accepted.

29. Find the probability that the claim will be rejected assuming that the manufacturer's claim is true.

30. Find the probability that the claim will be accepted assuming that the actual probability that the drug cures the skin disease is 65%.

Uses and Abuses

Statistics in the Real World

Uses

Normal Distributions Normal distributions can be used to describe many real-life situations and are widely used in the fields of science, business, and psychology. They are the most important probability distributions in statistics and can be used to approximate other distributions, such as discrete binomial distributions.

The most incredible application of the normal distribution lies in the Central Limit Theorem. This theorem states that no matter what type of distribution a population may have, as long as the sample size is at least 30, the distribution of sample means will be normal. If the population is itself normal, then the distribution of sample means will be normal no matter how small the sample is.

The normal distribution is essential to sampling theory. Sampling theory forms the basis of statistical inference, which you will begin to study in the next chapter.

Abuses

Confusing Likelihood with Certainty A common abuse of normal probability distributions is to confuse the concept of likelihood with the concept of certainty. For instance, if you randomly select a member from a population that is normally distributed, you know the probability is approximately 95% that you will obtain a value that lies within two standard deviations of the mean. This *does not* imply, however, that you cannot get an unusual result. In fact, 5% of the time you should expect to get a value that is more than two standard deviations from the mean.

Suppose a population is normally distributed with a mean of 100 and standard deviation of 15. It would not be unusual for an individual value taken from this population to be 112 or more. It *would* be, however, highly unusual to obtain a sample mean of 112 or more from a sample with 100 members.

Exercises

1. ***Confusing Likelihood with Certainty*** You are randomly selecting 100 people from a population that is normally distributed. Are you certain to get exactly 95 people who lie within two standard deviations of the mean? Explain your reasoning.
2. ***Confusing Likelihood with Certainty*** You are randomly selecting 10 people from a large population that is normally distributed. Which of the following is more likely? Explain your reasoning.
   a. All 10 lie within 2 standard deviations of the mean.
   b. At least one person does not lie within 2 standard deviations of the mean.

5 Chapter Summary

What did you learn? — Review Exercises

Section 5.1

What did you learn?	Review Exercises
◆ How to interpret graphs of normal probability distributions	*1, 2*
◆ How to find and interpret z-scores $z = \dfrac{x - \mu}{\sigma}$	*3, 4*
◆ How to find areas under the standard normal curve	*5–16*

Section 5.2

What did you learn?	Review Exercises
◆ How to find probabilities for normally distributed variables	*17–24*

Section 5.3

What did you learn?	Review Exercises
◆ How to find a z-score given the area under the normal curve	*25–30*
◆ How to transform a z-score to an x-value $x = \mu + z\sigma$	*31, 32*
◆ How to find a specific data value of a normal distribution given the probability	*33–36*

Section 5.4

What did you learn?	Review Exercises
◆ How to find sampling distributions and verify their properties	*37, 38*
◆ How to interpret the Central Limit Theorem $\mu_{\bar{x}} = \mu,\ \sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$	*39, 40*
◆ How to apply the Central Limit Theorem to find the probability of a sample mean	*41–46*

Section 5.5

What did you learn?	Review Exercises
◆ How to decide when the normal distribution can approximate the binomial distribution $\mu = np,\ \sigma = \sqrt{npq}$	*47, 48*
◆ How to find the correction for continuity	*49–52*
◆ How to use the normal distribution to approximate binomial probabilities	*53, 54*

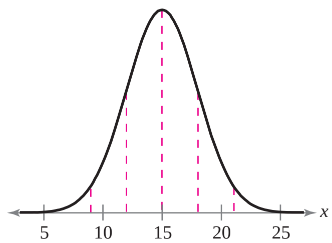
5 Review Exercises

1. $\mu = 15, \sigma = 3$
2. $\mu = -3, \sigma = 5$
3. $-2.25; 0.5; 2; 3.5$
4. 1.32 and 1.78 are unusual.
5. 0.2005
6. 0.9946
7. 0.3936
8. 0.8962
9. 0.0465
10. 0.7967
11. 0.4495
12. 0.2224
13. 0.3519
14. 0.95
15. 0.1336
16. 0.5905
17. 0.8997
18. 0.7704
19. 0.9236
20. 0.3364
21. 0.0124
22. 0.5465

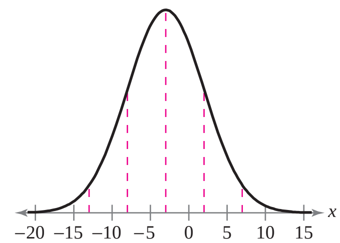
Section 5.1

In Exercises 1 and 2, use the graph to estimate μ and σ.

1.



2.



In Exercises 3 and 4, use the following information and standard scores to investigate observations about a normal population. A batch of 2500 resistors is normally distributed, with a mean resistance of 1.5 ohms and a standard deviation of 0.08 ohm. Four resistors are randomly selected and tested. Their resistances were measured at 1.32, 1.54, 1.66, and 1.78 ohms.

3. How many standard deviations from the mean are these observations?

4. Are there any unusual observations?

In Exercises 5–16, use the Standard Normal Table to find the indicated area under the standard normal curve.

5. To the left of $z = -0.84$

6. To the left of $z = 2.55$

7. To the left of $z = -0.27$

8. To the left of $z = 1.26$

9. To the right of $z = 1.68$

10. To the right of $z = -0.83$

11. Between $z = -1.64$ and the mean

12. Between $z = -1.22$ and $z = -0.43$

13. Between $z = 0.15$ and $z = 1.35$

14. Between $z = -1.96$ and $z = 1.96$

15. To the left of $z = -1.5$ and to the right of $z = 1.5$

16. To the left of $z = 0.12$ and to the right of $z = 1.72$

Section 5.2

In Exercises 17–22, find the indicated probabilities.

17. $P(z < 1.28)$

18. $P(z > -0.74)$

19. $P(-2.15 < z < 1.55)$

20. $P(0.42 < z < 3.15)$

21. $P(z < -2.50 \text{ or } z > 2.50)$

22. $P(z < 0 \text{ or } z > 1.68)$

23. (a) 0.3156
(b) 0.3099
(c) 0.3446
24. (a) 0.9544
(b) 0.3420
(c) 0.0026
25. −0.07
26. −1.28
27. 1.13
28. −2.055
29. 1.04
30. −0.84
31. 43.9 meters
32. 45.7 meters
33. 45.9 meters
34. 45.435 meters
35. 45.74 meters
36. 44.28 meters

In Exercises 23 and 24, find the indicated probabilities.

23. A study found that the mean migration distance of the green turtle was 2200 kilometers and the standard deviation was 625 kilometers. Assuming that the distances are normally distributed, find the probability that a randomly selected green turtle migrates a distance of

(a) less than 1900 kilometers.
(b) between 2000 kilometers and 2500 kilometers.
(c) greater than 2450 kilometers.

(Adapted from Dorling Kindersley Visual Encyclopedia)

24. The world's smallest mammal is the Kitti's hog-nosed bat, with a mean weight of 1.5 grams and a standard deviation of 0.25 gram. Assuming that the weights are normally distributed, find the probability of randomly selecting a bat that weighs

(a) between 1.0 gram and 2.0 grams.
(b) between 1.6 grams and 2.2 grams.
(c) more than 2.2 grams.

(Adapted from Dorling Kindersley Visual Encyclopedia)

Section 5.3

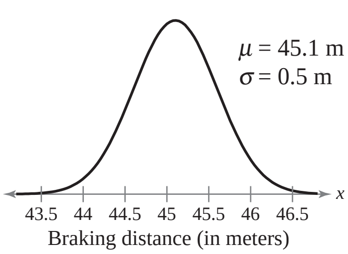
In Exercises 25–30, use the Standard Normal Table to find the z-score that corresponds to the given cumulative area or percentile. If the area is not in the table, use the entry closest to the area.

25. 0.4721 **26.** 0.1 **27.** 0.8708

28. P_2 **29.** P_{85} **30.** P_{20}

In Exercises 31–36, use the following information. On a dry surface, the braking distance (in meters) of a Pontiac Grand AM SE can be approximated by a normal distribution, as shown in the graph. *(Source: National Highway Traffic Safety Administration)*

Braking Distance of a Pontiac Grand Am SE



31. Find the braking distance of a Pontiac Grand AM SE that corresponds to $z = -2.4$.

32. Find the braking distance of a Pontiac Grand AM SE that corresponds to $z = 1.2$.

33. What braking distance of a Pontiac Grand AM SE represents the 95th percentile?

34. What braking distance of a Pontiac Grand AM SE represents the third quartile?

35. What is the shortest braking distance of a Pontiac Grand AM SE that can be in the top 10% of braking distances?

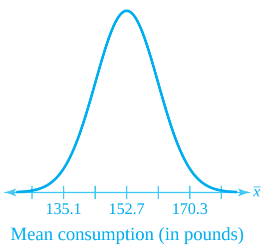
36. What is the longest braking distance of a Pontiac Grand AM SE that can be in the bottom 5% of braking distances?

37. See Odd Answers, page A##.

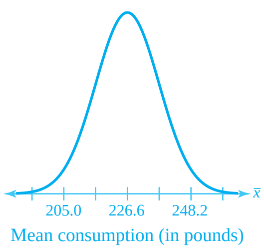
38. {00, 01, 02, 03, 10, 11, 12, 13, 20, 21, 22, 23, 30, 31, 32, 33}

1.5, 1.118; 1.5, 0.791

39. 152.7, 8.7



40. 226.6, 10.768



41. (a) 0.0485

(b) 0.8180

(c) 0.0823

(a) and (c) are smaller, (b) is larger. This is to be expected because the standard error of the sample means is smaller.

42. (a) ≈ 1

(b) 0.1446

(c) ≈ 0

(a) is larger and (b) and (c) are smaller.

43. (a) ≈ 0 (b) ≈ 0

44. (a) 0.9918

(b) 0.9998

Section 5.4

In Exercises 37 and 38, use the given population to find the sampling distribution of the sample means for the indicated sample sizes. Find the mean and standard deviation of the population and the mean and standard deviation of the sampling distribution. Compare the values.

37. A corporation has five executives. The number of minutes each exercises a week is reported as 40, 200, 80, 0, and 600. Draw three executives' names from this population, with replacement, and form a sampling distribution of the sample mean of the minutes they exercise.

38. There are four residents sharing a house. The number of times each washes his or her car each month is 1, 2, 0, and 3. Draw two names from this population, with replacement, and form a sampling distribution for the sample mean of the number of times their cars are washed each month.

In Exercises 39 and 40, use the Central Limit Theorem to find the mean and standard error of the mean of the indicated sampling distribution. Then sketch a graph of the sampling distribution.

39. The consumption of processed fruits by people in the United States in a recent year was normally distributed, with a mean of 152.7 pounds and a standard deviation of 51.6 pounds. Random samples of size 35 are drawn from this population. *(Adapted from U.S. Department of Agriculture)*

40. The consumption of processed vegetables by people in the United States in a recent year was normally distributed, with a mean of 226.6 pounds and a standard deviation of 68.1 pounds. Random samples of size 40 are drawn from this population. *(Adapted from U.S. Department of Agriculture)*

In Exercises 41–46, find the probabilities for the sampling distributions.

41. Refer to Exercise 23. A sample of 12 green turtles is randomly selected. Find the probability that the sample mean of the distance migrated is (a) less than 1900 kilometers, (b) between 2000 kilometers and 2500 kilometers, and (c) greater than 2450 kilometers. Compare your answers with those in Exercise 23.

42. Refer to Exercise 24. A sample of seven Kitti's hog-nosed bats is randomly selected. Find the probability that the sample mean is (a) between 1.0 gram and 2.0 grams, (b) between 1.6 grams and 2.2 grams, and (c) more than 2.2 grams. Compare your answers with those in Exercise 24.

43. The mean annual salary for chauffeurs is \$24,700. A random sample of size 45 is drawn from this population. What is the probability that the mean annual salary is (a) less than \$23,700 and (b) more than \$26,200? Assume $\sigma = \$1500$. *(Source: Salary.com)*

44. The mean value of land and buildings per acre for farms is \$1300. A random sample of size 36 is drawn. What is the probability that the mean value of land and buildings per acre is (a) less than \$1400 and (b) more than \$1150? Assume $\sigma = \$250$.

45. 0.0019

46. 0.0006

47. Cannot use normal distribution because $nq < 5$.

48. Can use normal distribution. $\mu = 8.85, \sigma = 1.905$

49. $P(x > 24.5)$

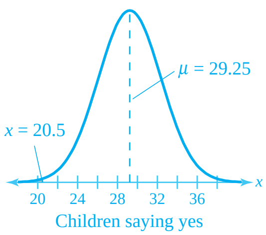
50. $P(x < 36.5)$

51. $P(44.5 < x < 45.5)$

52. $P(49.5 < x < 50.5)$

53. Can use normal distribution.

0.0032



54. Cannot use normal distribution because $np < 5$.

0.171

45. The mean price of houses in a city is \$1.5 million with a standard deviation of \$500,000. The house prices are normally distributed. You randomly select 15 houses in this city. What is the probability that the mean price will be less than \$1.125 million?

46. Mean rent in a city is \$5000 per month with a standard deviation of \$300. The rents are normally distributed. You randomly select 15 apartments in this city. What is the probability that the mean price will be more than \$5250?

Section 5.5

In Exercises 47 and 48, a binomial experiment is given. Decide whether you can use the normal distribution to approximate the binomial distribution. If you can, find the mean and standard deviation. If you cannot, explain why.

47. In a recent year, the American Cancer Society predicted that the five-year survival rate for new cases of kidney cancer would be 90%. You randomly select 12 men who were new kidney cancer cases this year and calculate their five-year survival rate. *(Source: American Cancer Society)*

48. A survey indicates that 59% of men purchased perfume in the past year. You randomly select 15 men and ask them if they have purchased perfume in the past year. *(Source: USA TODAY)*

In Exercises 49–52, write the binomial probability as a normal probability using the continuity correction.

	Binomial Probability	Normal Probability
49.	$P(x \geq 25)$	$P(x > ?)$
50.	$P(x \leq 36)$	$P(x < ?)$
51.	$P(x = 45)$	$P(? < x < ?)$
52.	$P(x = 50)$	$P(? < x < ?)$

In Exercises 53 and 54, decide whether you can use the normal distribution to approximate the binomial distribution. If you can, use the normal distribution to approximate the indicated probabilities and sketch their graphs. If you cannot, explain why and use the binomial distribution to find the indicated probabilities.

53. Sixty-five percent of children ages 12 to 17 keep at least part of their savings in a savings account. You randomly select 45 children and ask each if he or she keeps at least part of his or her savings in a savings account. Find the probability that at most 20 children will say yes. *(Source: International Communications Research for Merrill Lynch)*

54. Thirty-three percent of adults graded public schools as excellent or good at preparing students for college. You randomly select 12 adults and ask them if they think public schools are excellent or good at preparing students for college. Find the probability that more than five adults will say yes. *(Source: Marist Institute for Public Opinion)*

5 Chapter Quiz

1. (a) 0.9821
   (b) 0.9994
   (c) 0.9802
   (d) 0.8135
2. (a) 0.9198
   (b) 0.1940
   (c) 0.0456
3. 0.1611
4. 0.5739
5. 81.59%
6. 1417.6
7. 337.588
8. 257.952
9. ≈ 0
10. More likely to select one student with a test score greater than 300 because the standard error of the mean is less than the standard deviation.
11. Can use normal distribution. $\mu = 16.32, \sigma \approx 2.285$
12. 0.3594

Take this quiz as you would take a quiz in class. After you are done, check your work against the answers given in the back of the book.

1. Find each standard normal probability.

(a) $P(z > -2.10)$
(b) $P(z < 3.22)$
(c) $P(-2.33 < z < 2.33)$
(d) $P(z < -1.75 \text{ or } z > -0.75)$

2. Find each normal probability for the given parameters.

(a) $\mu = 5.5,\ \sigma = 0.08,\ P(5.36 < x < 5.64)$
(b) $\mu = -8.2,\ \sigma = 7.84,\ P(-5.00 < x < 0)$
(c) $\mu = 18.5,\ \sigma = 9.25,\ P(x < 0 \text{ or } x > 37)$

In Exercises 3–10, use the following information. In a recent year, grade 8 Washington State public school students taking a mathematics assessment test had a mean score of 281 with a standard deviation of 34.4. Possible test scores could range from 0 to 500. Assume that the scores are normally distributed. *(Source: National Center for Educational Statistics)*

3. Find the probability that a student had a score higher than 315.

4. Find the probability that a student had a score between 250 and 305.

5. What percent of the students had a test score that is greater than 250?

6. If 2000 students are randomly selected, how many would be expected to have a test score that is less than 300?

7. What is the lowest score that would still place a student in the top 5% of the scores?

8. What is the highest score that would still place a student in the bottom 25% of the scores?

9. A random sample of 60 students is drawn from this population. What is the probability that the mean test score is greater than 300?

10. Are you more likely to randomly select one student with a test score greater than 300 or are you more likely to select a sample of 15 students with a mean test score greater than 300? Explain.

In Exercises 11 and 12, use the following information. In a survey of adults, 68% thought that DNA tests for identifying an individual were very reliable. You randomly select 24 adults and ask each if he or she thinks DNA tests for identifying an individual are very reliable. *(Source: CBS News)*

11. Decide whether you can use the normal distribution to approximate the binomial distribution. If you can, find the mean and standard deviation. If you cannot, explain why.

12. Find the probability that at most 15 people say DNA tests for identifying an individual are very reliable.

PUTTING IT ALL TOGETHER

Real Statistics ▪ Real Decisions

You work for a manufacturing company as a statistical process analyst. Your job is to analyze processes and make sure they are in statistical control. In one process, a machine cuts wood boards to a thickness of 25 millimeters with an acceptable margin of error of ± 0.6 millimeter. (Assume this process can be approximated by a normal distribution.) So, the acceptable range of thicknesses for the boards is 24.4 millimeters to 25.6 millimeters, inclusive.

Because of machine vibrations and other factors, the setting of the wood-cutting machine "shifts" from 25 millimeters. To check that the machine is cutting the boards to the correct thickness, you select at random three samples of four boards and find the mean thickness (in millimeters) of each sample. A coworker asks you why you take three samples of size 4 and find the mean instead of randomly choosing and measuring 12 boards individually to check the machine's settings. (*Note:* Both samples are chosen without replacement.)

Exercises

1. ***Sampling Individuals***

You select one board and measure its thickness. Assume the machine shifts and is cutting boards with a mean thickness of 25.4 millimeters and a standard deviation of 0.2 millimeter.

(a) What is the probability that you select a board that is *not* outside the acceptable range (in other words, you do not detect that the machine has shifted)? (See figure.)

(b) You randomly select 12 boards. What is the probability that you select at least one board that is *not* outside the acceptable range?

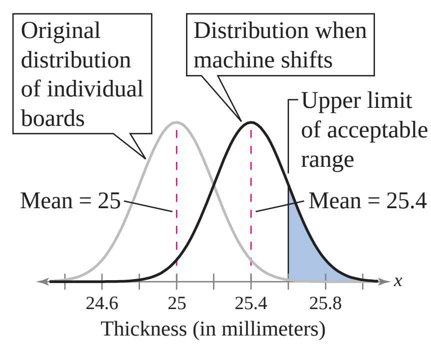


Figure for Exercise 1

2. ***Sampling Groups of Four***

You select four boards and find their mean thickness. Assume the machine shifts and is cutting boards with a mean thickness of 25.4 millimeters and a standard deviation of 0.2 millimeter.

(a) What is the probability that you select a sample of four boards that has a mean that is *not* outside the acceptable range? (See figure.)

(b) You randomly select three samples of four boards. What is the probability that you select at least one sample of four boards that has a mean that is *not* outside the acceptable range?

(c) What is more sensitive to change—an individual measure or the mean?

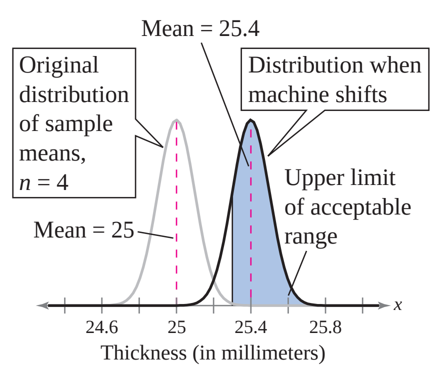


Figure for Exercise 2

3. ***Writing an Explanation***

Write a paragraph to your coworker explaining why you take three samples of size 4 and find the mean of each sample instead of randomly choosing and measuring 12 boards individually to check the machine's settings.

TECHNOLOGY | MINITAB | EXCEL | TI-83

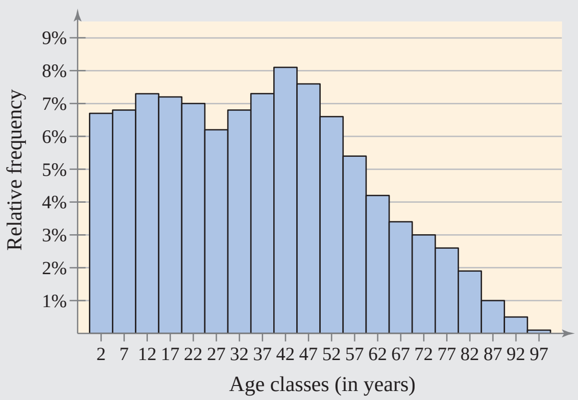
U.S. Census Bureau

www.census.gov

Age Distribution in the United States

One of the jobs of the U.S. Census Bureau is to keep track of the age distribution in the country. The age distribution in 2003 is shown below.

Age Distribution in the U.S.



Class boundaries	Class midpoint	Relative frequency
0–4	2	6.7%
5–9	7	6.8%
10–14	12	7.4%
15–19	17	7.2%
20–24	22	7.0%
25–29	27	6.2%
30–34	32	6.8%
35–39	37	7.3%
40–44	42	8.1%
45–49	47	7.6%
50–54	52	6.6%
55–59	57	5.5%
60–64	62	4.2%
65–69	67	3.4%
70–74	72	3.0%
75–79	77	2.6%
80–84	82	1.9%
85–89	87	1.0%
90–94	92	0.5%
95–99	97	0.2%

Exercises

We used a technology tool to select random samples with $n = 40$ from the age distribution of the United States. The means of the 36 samples were as follows.

DATA

28.14, 31.56, 36.86, 32.37, 36.12, 39.53,
36.19, 39.02, 35.62, 36.30, 34.38, 32.98,
36.41, 30.24, 34.19, 44.72, 38.84, 42.87,
38.90, 34.71, 34.13, 38.25, 38.04, 34.07,
39.74, 40.91, 42.63, 35.29, 35.91, 34.36,
36.51, 36.47, 32.88, 37.33, 31.27, 35.80

1. Enter the age distribution of the United States into a technology tool. Use the tool to find the mean age in the United States.

2. Enter the set of sample means into a technology tool. Find the mean of the set of sample means. How does it compare with the mean age in the United States? Does this agree with the result predicted by the Central Limit Theorem?

3. Are the ages of people in the United States normally distributed? Explain your reasoning.

4. Sketch a relative frequency histogram for the 36 sample means. Use nine classes. Is the histogram approximately bell shaped and symmetric? Does this agree with the result predicted by the Central Limit Theorem?

5. Use a technology tool to find the standard deviation of the ages of people in the United States.

6. Use a technology tool to find the standard deviation of the set of 36 sample means. How does it compare with the standard deviation of the ages? Does this agree with the result predicted by the Central Limit Theorem?

Extended solutions are given in the *Technology Supplement.* Technical instruction is provided for MINITAB, Excel, and the TI-83.

d.

x	$P(x)$
0	0.0546
1	0.197
2	0.304
3	0.261
4	0.134
5	0.0416
6	0.00714
7	0.000525
	$\Sigma P(x) \approx 1$

4a. $n = 250$, $p = 0.71$, $x = 178$ **b.** 0.056

c. The probability that exactly 178 people in the United States will use more than one topping on their hotdog is about 0.056.

5a. (1) $x = 2$ (2) $x = 2, 3, 4$, or 5 (3) $x = 0$ or 1

b. (1) 0.217

(2) 0.217, 0.058, 0.008, 0.0004; 0.283

(3) 0.308, 0.409; 0.717

c. (1) The probability that exactly two men consider fishing their favorite leisure-time activity is about 0.217.

(2) The probability that at least two men consider fishing their favorite leisure-time activity is about 0.283.

(3) The probability that fewer than two men consider fishing their favorite leisure-time activity is about 0.717.

6a. Trial: selecting a business and asking if it has a website

Success: selecting a business with a website

Failure: selecting a business without a website

b. $n = 10$, $p = 0.39$, $x = 4$ **c.** 0.200

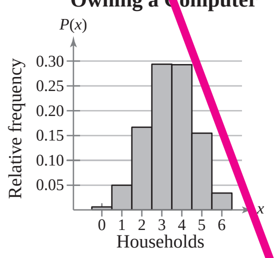
d. The probability that exactly four of the 10 small business have a website is 0.200.

7a. 0.006, 0.050, 0.167, 0.294, 0.293, 0.155, 0.034

b.

x	$P(x)$
0	0.006
1	0.050
2	0.167
3	0.294
4	0.293
5	0.155
6	0.034

c.



8a. Success: selecting a clear day

$n = 31$, $p = 0.44$, $q = 0.56$

b. 13.6 **c.** 7.6 **d.** 2.8

e. On average, there are about 14 clear days during the month of May. The standard deviation is about 3 days.

Section 4.3

1a. 0.23, 0.177, 0.136 **b.** 0.543

c. The probability that your first sale will occur before your fourth sales call is 0.543.

2a. $P(0) \approx 0.050$

$P(1) \approx 0.149$

$P(2) \approx 0.224$

$P(3) \approx 0.224$

$P(4) \approx 0.168$

b. 0.815 **c.** 0.185

d. The probability that more than four accidents will occur in any given month at the intersection is 0.185.

3a. 0.10 **b.** 0.10, 3 **c.** 0.0002

d. The probability of finding three brown trout in any given cubic meter of the lake is 0.0002.

CHAPTER 5

Section 5.1

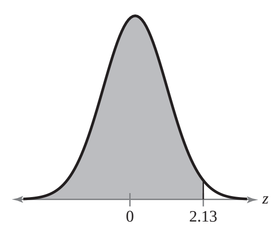
1a. A: 45, B: 60, C: 45; B has the greatest mean.

b. Curve C is more spread out, so curve C has the greatest standard deviation.

2a. 3.5 feet **b.** 3.3, 3.7; 0.2 foot

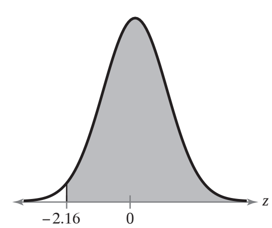
3a. (1) 0.0143 (2) 0.9850

4a.



b. 0.9834

5a.



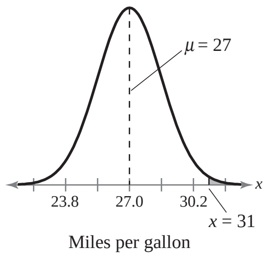
b. 0.0154

c. 0.9846

6a. 0.0885 **b.** 0.0154 **c.** 0.0731

Section 5.2

1a.

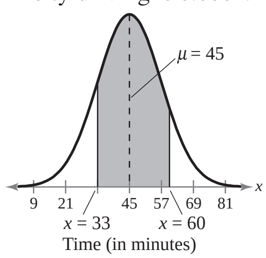


b. 2.50

c. 0.0062

d. The probability that a randomly selected manual transmission Focus will get more than 31 miles per gallon in city driving is 0.0062.

2a.



b. -1, 1.25

c. 0.1587; 0.8944

d. 0.7357

3a. Read user's guide for the technology tool.

b. Enter the data.

c. The probability that a randomly selected U.S. man's cholesterol is between 190 and 225 is about 0.4968.

Section 5.3

1a. (1) 0.0384 (2) 0.0250 and 0.9750

bc. (1) -1.77 (2) ± 1.96

2a. (1) Area $= 0.10$ (2) Area $= 0.20$
(3) Area $= 0.99$

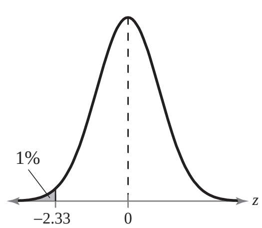
bc. (1) -1.28 (2) -0.84 (3) 2.33

3a. $\mu = 70$, $\sigma = 8$

b. 64; 104.32; 55.44

c. 64 is below the mean, 104.32 is above the mean, and 55.44 is below the mean.

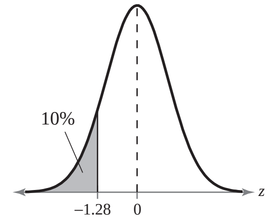
4ab.



c. 142.83

d. So, the longest braking distance a Ford F-150 could have and still be in the top 1% is 143 feet.

5ab.



c. 8.512

d. So, the maximum length of time an employee could have worked and still be laid off is 8 years.

Section 5.4

1a.

Sample	Mean	Sample	Mean	Sample	Mean
1, 1, 1	1	3, 3, 5	3.67	5, 7, 1	4.33
1, 1, 3	1.67	3, 3, 7	4.33	5, 7, 3	5
1, 1, 5	2.33	3, 5, 1	3	5, 7, 5	5.67
1, 1, 7	3	3, 5, 3	3.67	5, 7, 7	6.33
1, 3, 1	1.67	3, 5, 5	4.33	7, 1, 1	3
1, 3, 3	2.33	3, 5, 7	5	7, 1, 3	3.67
1, 3, 5	3	3, 7, 1	3.67	7, 1, 5	4.33
1, 3, 7	3.67	3, 7, 3	4.33	7, 1, 7	5
1, 5, 1	2.33	3, 7, 5	5	7, 3, 1	3.67
1, 5, 3	3	3, 7, 7	5.67	7, 3, 3	4.33
1, 5, 5	3.67	5, 1, 1	2.33	7, 3, 5	5
1, 5, 7	4.33	5, 1, 3	3	7, 3, 7	5.67
1, 7, 1	3	5, 1, 5	3.67	7, 5, 1	4.33
1, 7, 3	3.67	5, 1, 7	4.33	7, 5, 3	5
1, 7, 5	4.33	5, 3, 1	3	7, 5, 5	5.67
1, 7, 7	5	5, 3, 3	3.67	7, 5, 7	6.33
3, 1, 1	1.67	5, 3, 5	4.33	7, 7, 1	5
3, 1, 3	2.33	5, 3, 7	5	7, 7, 3	5.67
3, 1, 5	3	5, 5, 1	3.67	7, 7, 5	6.33
3, 1, 7	3.67	5, 5, 3	4.33	7, 7, 7	7
3, 3, 1	2.33	5, 5, 5	5		
3, 3, 3	3	5, 5, 7	5.67		

b.

$\bar{x}$	f	Probability
1	1	0.0156
1.67	3	0.0469
2.33	6	0.0938
3	10	0.1563
3.67	12	0.1875
4.33	12	0.1875
5	10	0.1563
5.67	6	0.0938
6.33	3	0.0469
7	1	0.0156

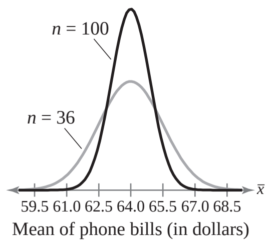
$$\mu_{\bar{x}} = 4$$
$$(\sigma_{\bar{x}})^2 = 1.667$$
$$\sigma_{\bar{x}} = 1.291$$

c. $\mu_{\bar{x}} = \mu = 4$

$$(\sigma_{\bar{x}})^2 = \frac{\sigma^2}{n} = \frac{\sigma}{3} = 1.667; \sigma_{\bar{x}} = \frac{\sigma}{n} = \frac{\sqrt{5}}{\sqrt{3}} = 1.291$$

2a. 64, 0.9

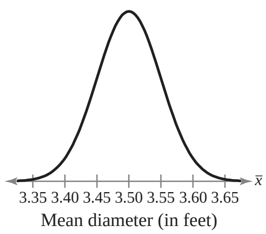
b. $n = 100$



c. With a larger sample size, the standard deviation decreases.

3a. 3.5, 0.05

b.

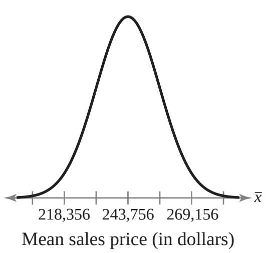


4a. $\mu_{\bar{x}} = 25$

$\sigma_{\bar{x}} = 0.15$

b. −2, 3.33 **c.** 0.9768

5a. 243,756; 12,701.71



b. −3.44 **c.** 0.9997

6a. 0.5, 1.58 **b.** 0.6915, 0.9429

c. There is a 69% chance an *individual receiver* will cost less than $700. There is a 94% chance that the *mean of a sample of 10 receivers* is less than $700.

Section 5.5

1a. 70, 0.61, 0.39 **b.** 42.7, 27.3

c. Normal distribution can be used.

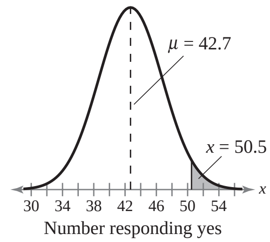
d. 42.7, 4.08

2a. (1) 57, 58, . . . , 83 (2) . . . , 52, 53, 54

b. (1) $56.5 < x < 83.5$ (2) $x < 54.5$

3a. Normal distribution can be used. **b.** 42.7, 4.08

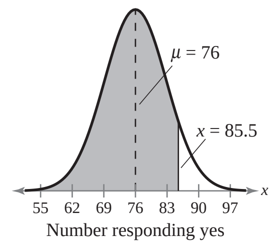
c.



d. 1.91 **e.** 0.0281

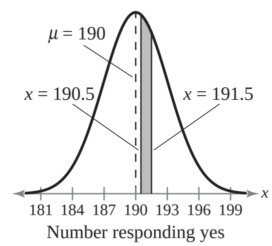
4a. Normal distribution can be used. **b.** 76, 6.86

c. $P(x \leq 85.5)$ **d.** 1.38 **e.** 0.9162



5a. Normal distribution can be used. **b.** 190, 3.08

c. $P(190.5 < x < 191.5)$ **d.** 0.16, 0.49 **e.** 0.1243



CHAPTER 6

Real Statistics–Real Decisions for Chapter 4

(page 212)

1. (a) Answers will vary. For example, calculate the probability of obtaining zero clinical pregnancies out of 10 randomly selected ART cycles.
   (b) Binomial. The distribution is discrete because the number of clinical pregnancies is countable.
2. $n = 10$, $p = 0.328$, $P(0) = 0.0188$

x	$P(x)$
0	0.0188
1	0.0917
2	0.2013
3	0.2621
4	0.2238
5	0.1311
6	0.0533
7	0.0149
8	0.0027
9	0.0003
10	0.00001

3. (a) Suspicious, because the probability is very small.
   (b) Not suspicious, because the probability is not that small.

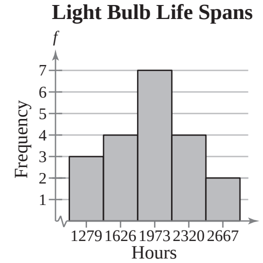
CHAPTER 5

Section 5.1 *(page 224)*

1. Answers will vary.
3. Answers will vary.
   Similarities: The two curves will have the same line of symmetry.
   Differences: One curve will be more spread out than the other.
5. $\mu = 0, \sigma = 1$
7. "The" standard normal distribution is used to describe one specific normal distribution $(\mu = 0, \sigma = 1)$. "A" normal distribution is used to describe a normal distribution with any mean and standard deviation.
9. No, the graph crosses the x-axis.
11. Yes, the graph fulfills the properties of the normal distribution.
13. No, the graph is skewed right.
15. It is normal because it is bell shaped and symmetric.

17. 0.3849 **19.** 0.6247 **21.** 0.9382 **23.** 0.975
25. 0.8289 **27.** 0.1003 **29.** 0.005 **31.** 0.05
33. 0.475 **35.** 0.437 **37.** 0.95 **39.** 0.2006

41. (a)



It is reasonable to assume that the life span is normally distributed because the histogram is nearly symmetric and bell shaped.

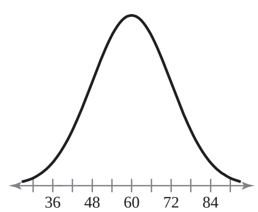
(b) 1941.35, 432.385
(c) The sample mean of 1941.35 hours is less than the claimed mean, so, on average, the bulbs in the sample lasted for a shorter time. The sample standard deviation of 432 hours is greater than the claimed standard deviation, so the bulbs in the sample had a greater variation in life span than the manufacturer's claim.

43. (a) A = 2.97; B = 2.98; C = 3.01; D = 3.05
(b) 0.5; −1.5; −1; 2.5
(c) $x = 3.05$ is unusual owing to a relatively large z-score (2.5).

45. (a) A = 801; B = 950; C = 1250; D = 1467
(b) −0.36; 1.07; 2.11; −1.08
(c) $x = 1467$ is unusual owing to a relatively large z-score (2.11).

47. 0.6915 **49.** 0.05 **51.** 0.5328 **53.** 0.9265
55. 0.9744 **57.** 0.3133 **59.** 0.901 **61.** 0.0098

63.



The normal distribution curve is centered at its mean (60) and has 2 points of inflection (48 and 72) representing $\mu \pm \sigma$.

65. (a) Area under curve = area of rectangle
$= (1)(1) = 1$
(b) 0.25 (c) 0.4

Section 5.2 *(page 232)*

1. 0.1151 **3.** 0.1151 **5.** 0.1144
7. 0.3022 **9.** 0.2742 **11.** 0.0566
13. (a) 0.1357 (b) 0.6983 (c) 0.1660
15. (a) 0.1711 (b) 0.7018 (c) 0.1271
17. (a) 0.0062 (b) 0.9876 (c) 0.0062
19. (a) 0.0073 (b) 0.806 (c) 0.1867
21. (a) 79.95% (b) 348 **23.** (a) 64.8% (b) 18
25. (a) 30.85% (b) 31 **27.** (a) 99.87% (b) 0.798
29. 1.5%; It is unusual for a battery to have a life span that is more than 2065 hours because of the relatively large z-score (2.17).
31. Out of control, because there is a point more than 3 standard deviations beyond the mean.

33. Out of control, because there are nine consecutive points below the mean, and two out of three consecutive points lie more than 2 standard deviations from the mean.

Section 5.3 *(page 242)*

1. −2.05 **3.** 0.85 **5.** −0.16 **7.** 2.39
9. −1.645 **11.** 0.995 **13.** −2.325 **15.** −0.25
17. 1.175 **19.** −0.675 **21.** 0.675 **23.** −0.385
25. −0.38 **27.** −0.58 **29.** −1.645, 1.645
31. 0.325 **33.** −0.33 **35.** 1.28 **37.** ±0.06
39. (a) 68.52 (b) 62.14 **41.** (a) 12.28 (b) 20.08
43. (a) 139.22 (b) 96.92 **45.** 19.89
47. Tires that wear out by 26,800 miles will be replaced free of charge.
49. 7.93

Section 5.4 *(page 254)*

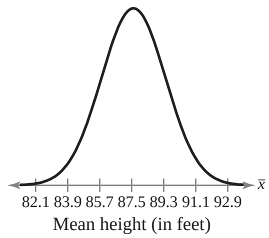
1. 100, 2.12 **3.** 100, 0.949
5. False. As the size of a sample increases, the mean of the distribution of sample means does not change.
7. False. The shape of a sampling distribution is normal if either $n \geq 30$ or the shape of the population is normal.
9.

Sample	Mean	Sample	Mean	Sample	Mean
0, 0, 0	0	2, 2, 4	2.67	4, 8, 0	4
0, 0, 2	0.67	2, 2, 8	4	4, 8, 2	4.67
0, 0, 4	1.33	2, 4, 0	2	4, 8, 4	5.33
0, 0, 8	2.67	2, 4, 2	2.67	4, 8, 8	6.67
0, 2, 0	0.67	2, 4, 4	3.33	8, 0, 0	2.67
0, 2, 2	1.33	2, 4, 8	4.67	8, 0, 2	3.33
0, 2, 4	2	2, 8, 0	3.33	8, 0, 4	4
0, 2, 8	3.33	2, 8, 2	4	8, 0, 8	5.33
0, 4, 0	1.33	2, 8, 4	4.67	8, 2, 0	3.33
0, 4, 2	2	2, 8, 8	6	8, 2, 2	4
0, 4, 4	2.67	4, 0, 0	1.33	8, 2, 4	4.67
0, 4, 8	4	4, 0, 2	2	8, 2, 8	6
0, 8, 0	2.67	4, 0, 4	2.67	8, 4, 0	4
0, 8, 2	3.33	4, 0, 8	4	8, 4, 2	4.67
0, 8, 4	4	4, 2, 0	2	8, 4, 4	5.33
0, 8, 8	5.33	4, 2, 2	2.67	8, 4, 8	6.67
2, 0, 0	0.67	4, 2, 4	3.33	8, 8, 0	5.33
2, 0, 2	1.33	4, 2, 8	4.67	8, 8, 2	6
2, 0, 4	2	4, 4, 0	2.67	8, 8, 4	6.67
2, 0, 8	3.33	4, 4, 2	3.33	8, 8, 8	8
2, 2, 0	1.33	4, 4, 4	4		
2, 2, 2	2	4, 4, 8	5.33		

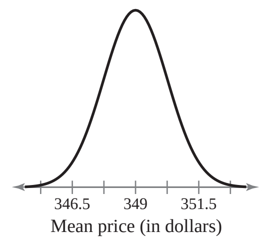
$\mu_{\bar{x}} = 3.5, \sigma_{\bar{x}} = 1.708$
$\mu = 3.5, \sigma = 2.958$

11. (c), because $\mu = 16.5$, $\sigma = 1.19$, and the graph approximates a normal curve.

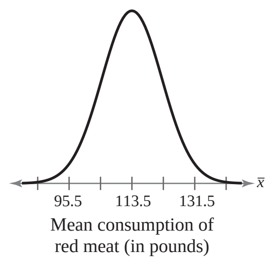
13. 87.5, 1.804



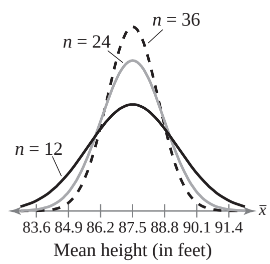
15. 349, 1.26



17. 113.5, 8.61



19. 87.5, 1.276; 87.5, 1.042



As the sample size increases, the standard error decreases, while the mean of the sample means remains constant.

21. 0.0019 **23.** 0.6319 **25.** ≈ 0
27. It is more likely to select a sample of 20 women with a mean height less than 70 inches because the sample of 20 has a higher probability.
29. Yes, it is very unlikely that you would have randomly sampled 40 cans with a mean equal to 127.9 ounces.
31. (a) 0.0008 (b) Claim is inaccurate.
(c) No, assuming the manufacturer's claim is true, because 96.25 is within 1 standard deviation of the mean for an individual board.
33. (a) 0.0002 (b) Claim is inaccurate.
(c) No, assuming the manufacturer's claim is true, because 49,721 is within 1 standard deviation of the mean for an individual tire.
35. Yes, because of the relatively large z-score (2.12).
37. ≈ 0

Section 5.5 *(page 265)*

1. Cannot use normal distribution.
3. Can use normal distribution.
5. Cannot use normal distribution because $np < 5$.

7. Cannot use normal distribution because $nq < 5$.

9. d **10.** b **11.** a **12.** c **13.** a

14. d **15.** c **16.** b

17. Binomial: 0.549; Normal: 0.5463

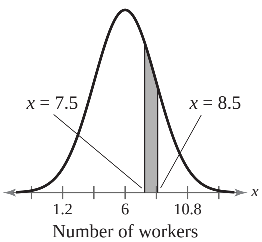
19. Cannot use normal distribution because $np < 5$.

(a) 0.0000199 (b) 0.000023

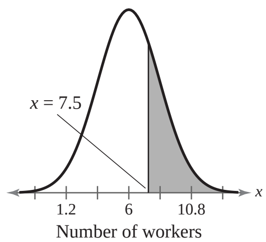
(c) 0.999977 (d) 0.1635

21. Can use normal distribution.

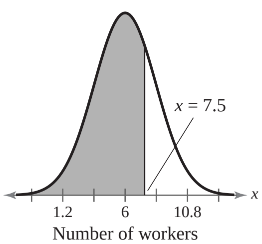
(a) 0.1174



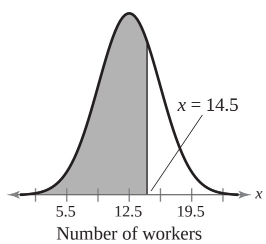
(b) 0.2643



(c) 0.7357

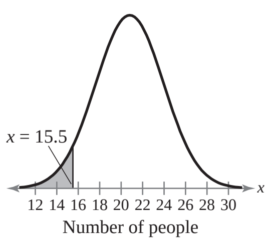


(d) 0.7190

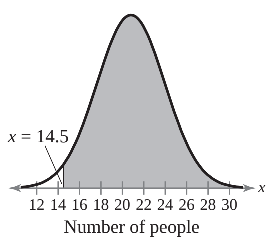


23. Can use normal distribution.

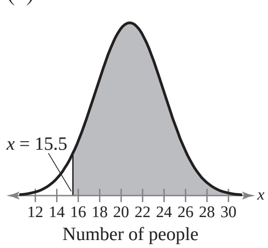
(a) 0.0465



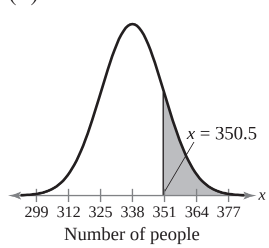
(b) 0.9767



(c) 0.9535



(d) 0.1635



25. (a) $np = 6 \geq 5$

$nq = 19 \geq 5$

(b) 0.121

(c) No, because the z-score is within one standard deviation of the mean.

27. Highly unlikely. Answers will vary.

29. 0.1020

Uses and Abuses for Chapter 5 *(page 269)*

1. No. Answers will vary.

2. It is more likely that all 10 people lie within 2 standard deviations of the mean. This can be shown by using the Empirical Rule and the Multiplication Rule.

(a) By the Empirical Rule, the probability of lying within 2 standard deviations of the mean is 0.95. Let x = number of people selected who lie within 2 standard deviations of the mean.

$P(x = 10) = (0.95)^{10} \approx 0.599$

(b) P (at least one person does not lie within 2 standard deviations of the mean) $= 1 - P(x = 10) \approx 1 - 0.599 = 0.401$

Review Answers for Chapter 5 *(page 271)*

1. $\mu = 15, \sigma = 3$ **3.** $-2.25; 0.5; 2; 3.5$ **5.** 0.2005

7. 0.3936 **9.** 0.0465 **11.** 0.4495 **13.** 0.3519

15. 0.1336 **17.** 0.8997 **19.** 0.9236 **21.** 0.0124

23. (a) 0.3156 (b) 0.3099 (c) 0.3446

25. -0.07 **27.** 1.13 **29.** 1.04 **31.** 43.9 meters

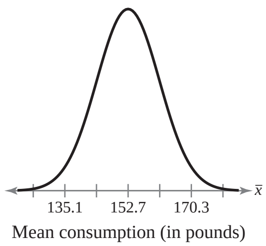
33. 45.9 meters **35.** 45.74 meters

37. {0 0 0, 0 0 200, 0 0 40, 0 0 600, 0 0 80, 0 200 0, 0 200 200, 0 200 40, 0 200 600, 0 200 80, 0 40 0, 0 40 200, 0 40 40, 0 40 600, 0 40 80, 0 600 0, 0 600 200, 0 600 40, 0 600 600, 0 600 80, 0 80 0, 0 80 200, 0 80 40, 0 80 600, 0 80 80, 200 0 0, 200 0 200, 200 0 40, 200 0 600, 200 0 80, 200 200 0, 200 200 200, 200 200 40, 200 200 600, 200 200 80, 200 40 0, 200 40 200, 200 40 40, 200 40 600, 200 40 80, 200 600 0, 200 600 200, 200 600 40, 200 600 600, 200 600 80, 200 80 0, 200 80 200, 200 80 40, 200 80 600, 200 80 80, 40 0 0, 40 0 200, 40 0 40, 40 0 600, 40 0 80, 40 200 0, 40 200 200, 40 200 40, 40 200 600, 40 200 80, 40 40 0, 40 40 200, 40 40 40, 40 40 600, 40 40 80, 40 600 0, 40 600 200, 40 600 40, 40 600 600, 40 600 80, 40 80 0, 40 80 200, 40 80 40, 40 80 600, 40 80 80, 600 0 0, 600 0 200, 600 0 40, 600 0 600, 600 0 80, 600 200 0, 600 200 200, 600 200 40, 600 200 600, 600 200 80, 600 40 0, 600 40 200, 600 40 40, 600 40 600, 600 40 80, 600 600 0, 600 600 200, 600 600 40, 600 600 600, 600 600 80, 600 80 0, 600 80 200, 600 80 40, 600 80 600, 600 80 80, 80 0 0, 80 0 200, 80 0 40, 80 0 600, 80 0 80, 80 200 0, 80 200 200, 80 200 40, 80 200 600, 80 200 80, 80 40 0, 80 40 200, 80 40 40, 80 40 600, 80 40 80, 80 600 0, 80 600 200, 80 600 40, 80 600 600, 80 600 80, 80 80 0, 80 80 200, 80 80 40, 80 80 600, 80 80 80}

$\mu = 184, \sigma \approx 218.504$

$\mu_{\bar{x}} = 184, \sigma_{\bar{x}} \approx 126.153$

39. 152.7, 8.7



41. (a) 0.0485 (b) 0.8180 (c) 0.0823

(a) and (c) are smaller, (b) is larger. This is to be expected because the standard error of the sample means is smaller.

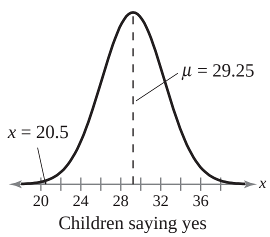
43. (a) ≈ 0 (b) ≈ 0 **45.** 0.0019

47. Cannot use normal distribution because $nq < 5$.

49. $P(x > 24.5)$ **51.** $P(44.5 < x < 45.5)$

53. Can use normal distribution.

0.0032



Chapter Quiz for Chapter 5 *(page 275)*

1. (a) 0.9821 (b) 0.9994 (c) 0.9802 (d) 0.8135

2. (a) 0.9198 (b) 0.1940 (c) 0.0456

3. 0.1611 **4.** 0.5739 **5.** 81.59% **6.** 1417.6

7. 337.588 **8.** 257.952 **9.** ≈ 0

10. More likely to select one student with a test score greater than 300 because the standard error of the mean is less than the standard deviation.

11. Can use normal distribution.

$\mu = 16.32, \sigma \approx 2.285$

12. 0.3594

Real Statistics–Real Decisions for Chapter 5

(page 276)

1. (a) 0.8413 (b) 0.9999999997

2. (a) 0.9772 (b) 0.9999881476 (c) mean

3. Answers will vary.

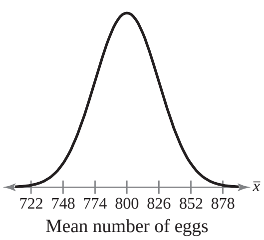
CHAPTER 6

CHAPTER 5

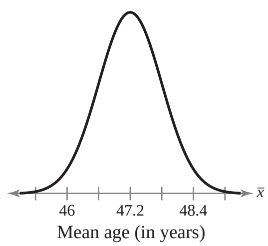
Section 5.4

12. (b), because $\mu = 5.8$, $\sigma = 0.23$, and the graph approximates a normal curve.

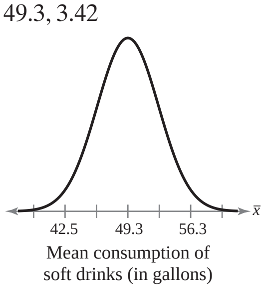
14. 800, 25.820



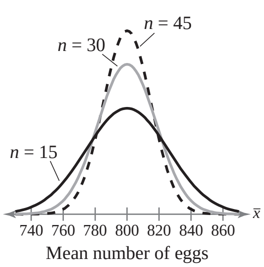
16. 47.2, 0.6



18. 49.3, 3.42



20. 800, 18.257; 800, 14.907

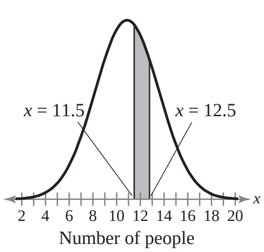


As the sample size increases, the standard error decreases.

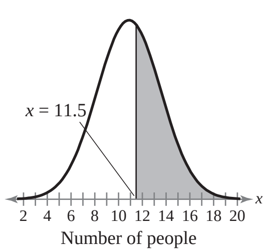
Section 5.5

20. Can use normal distribution.

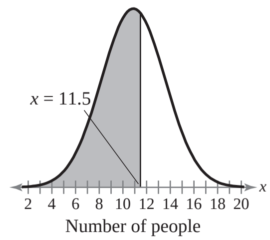
(a) 0.1347



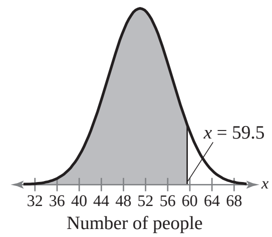
(b) 0.4090



(c) 0.5910

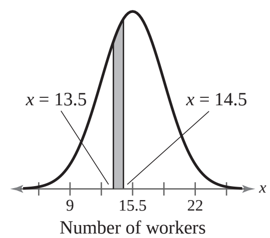


(d) 0.9292

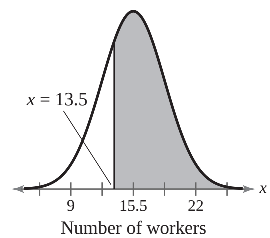


22. Can use normal distribution.

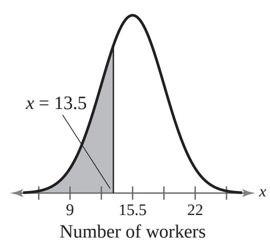
(a) 0.1074



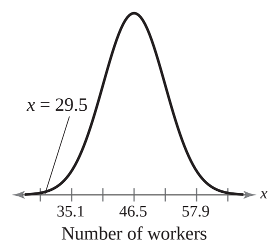
(b) 0.7291



(c) 0.2709



(d) 0.0013



CHAPTER 6