5

CHAPTER

NORMAL PROBABILITY DISTRIBUTIONS

The bottom shell of an Eastern Box Turtle has hinges so the turtle can retract its head, tail, and legs into the shell. The shell can also regenerate if it has been damaged.

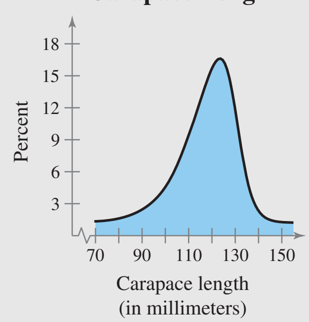
WHERE YOU'VE BEEN

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 determine the rate of clinical mastitis (infections caused by bacteria that can alter milk production) in dairy herds. How would you organize the study? When the Animal Health Service performed this study, it used random sampling and then classified the results according to breed, housing, hygiene, health, milking management, and milking machine. One conclusion from the study was that herds with Red and White cows as the predominant breed had a higher rate of clinical mastitis than herds with Holstein-Friesian cows as the main breed.

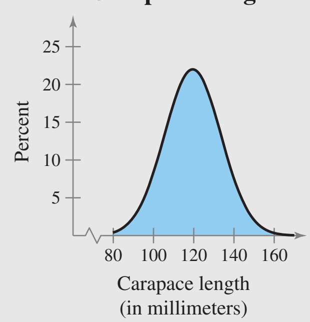
WHERE YOU'RE GOING

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 North Carolina Zoo and were collecting data about various physical traits of Eastern Box Turtles at the zoo. Which of the following would you expect to have bell-shaped, symmetric distributions: carapace (top shell) length, plastral (bottom shell) length, carapace width, plastral width, weight, total length? For instance, the four graphs below show the carapace length and plastral length of male and female Eastern Box Turtles in the North Carolina Zoo. Notice that the male Eastern Box Turtle carapace length distribution is bell-shaped, but the other three distributions are skewed left.

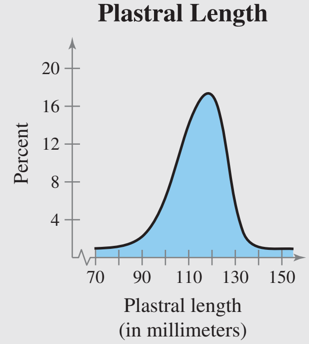
Female Eastern Box Turtle Carapace Length



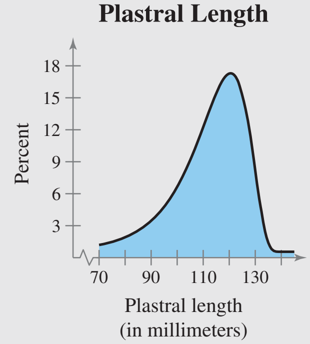
Male Eastern Box Turtle Carapace Length



Female Eastern Box Turtle Plastral Length



Male Eastern Box Turtle Plastral Length



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 distinguished between discrete and continuous random variables, and 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 pressures of humans, the lifetimes of plasma televisions, and even housing costs are all normally distributed random variables.

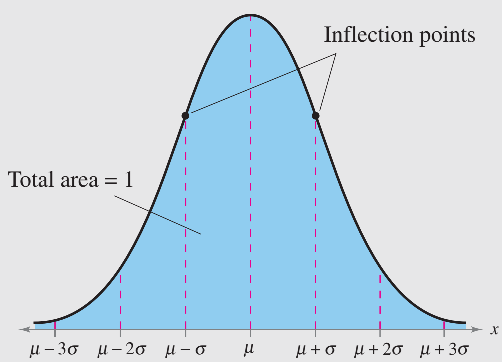
DEFINITION

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 1.
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

To learn how to determine if a random sample is taken from a normal distribution, see Appendix C.



INSIGHT

A probability density function has two requirements.

1. The total area under the curve is equal to 1.
2. The function can never be negative.

You have learned that a discrete probability distribution can be graphed with a histogram. For a continuous probability distribution, you can use a **probability density function (pdf).** A normal curve with mean μ and standard deviation σ can be graphed using the normal probability density function.

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

A normal curve depends completely on the two parameters μ and σ because $e \approx 2.718$ and $\pi \approx 3.14$ are constants.

STUDY TIP

Here are instructions for graphing a normal distribution on a TI-83/84 Plus.

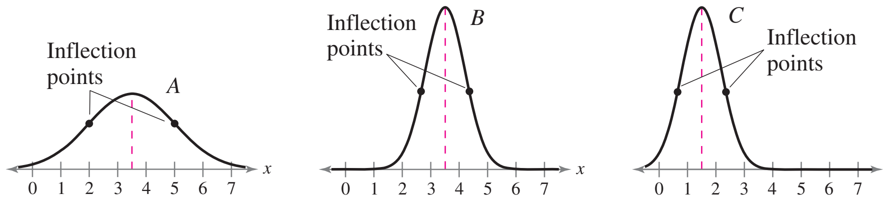
Y= 2nd DISTR

1: normalpdf(

Enter x and the values of μ and σ separated by commas.

GRAPH

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: $\sigma = 0.7$

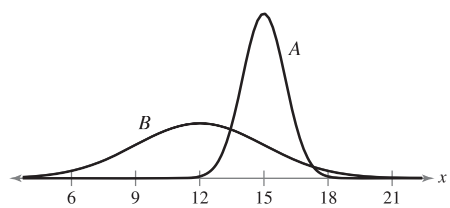
Mean: $\mu = 1.5$
Standard deviation: $\sigma = 0.7$

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?

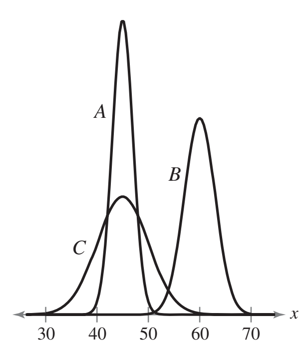


Solution

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

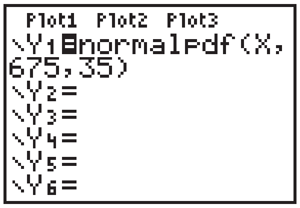
Consider the normal curves shown at the right. Which normal curve has the greatest mean? Which normal curve has the greatest standard deviation?

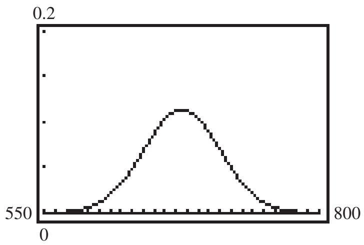


a. Find the location of the *line of symmetry* of each curve. Make a conclusion about which mean is greatest.

b. Determine which normal curve is *more spread out*. Make a conclusion about which standard deviation is greatest.

Answer: Page A37



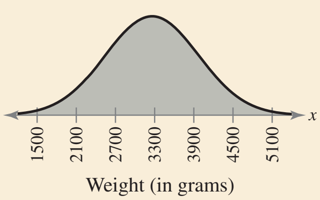


Once you determine the mean and standard deviation, you can use a TI-83/84 Plus to graph the normal curve in Example 2.

PICTURING THE WORLD

According to one publication, the number of births in the United States in a recent year was 4,317,000. The weights of the newborns can be approximated by a normal distribution, as shown by the following graph. (Adapted from National Center for Health Statistics)

Weights of Newborns

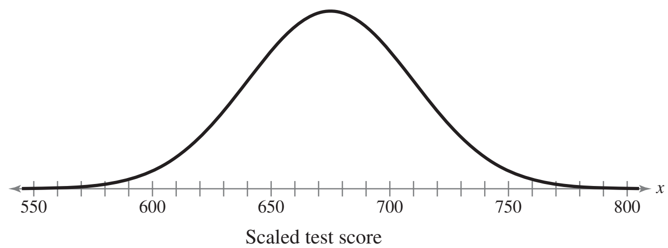


What is the mean weight of the newborns? Estimate the standard deviation of this normal distribution.

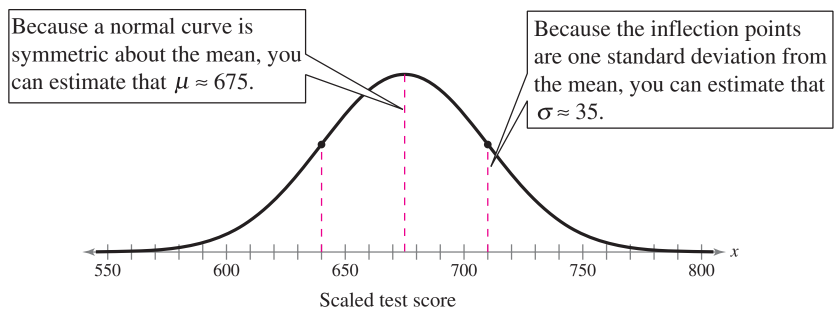
EXAMPLE 2

Interpreting Graphs of Normal Distributions

The scaled test scores for the New York State Grade 8 Mathematics Test are normally distributed. The normal curve shown below represents this distribution. What is the mean test score? Estimate the standard deviation of this normal distribution. *(Adapted from New York State Education Department)*



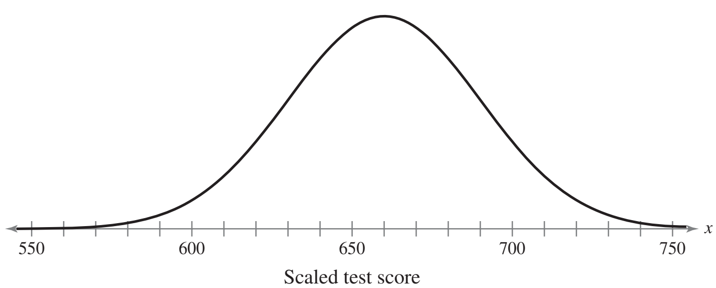
Solution



Interpretation The scaled test scores for the New York State Grade 8 Mathematics Test are normally distributed with a mean of about 675 and a standard deviation of about 35.

Try It Yourself 2

The scaled test scores for the New York State Grade 8 English Language Arts Test are normally distributed. The normal curve shown below represents this distribution. What is the mean test score? Estimate the standard deviation of this normal distribution. *(Adapted from New York State Education Department)*



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

Answer: Page A37

▶ 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

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

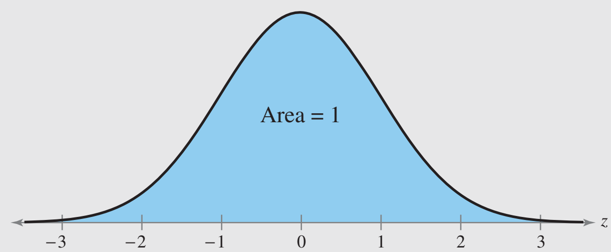
$$= \frac{x - \mu}{\sigma}.$$ Round to the nearest hundredth.

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 probabilities) 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

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 $z = 3.49$.

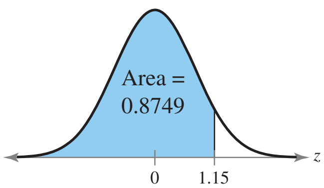
EXAMPLE 3

Using the Standard Normal Table

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

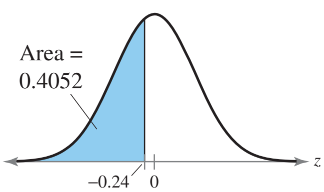
2. Find the cumulative area that corresponds to a z-score of -0.24.

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

You can also use a computer or calculator to find the cumulative area that corresponds to a z-score, as shown in the margin.

STUDY TIP

Here are instructions for finding the area that corresponds to $z = -0.24$ on a TI-83/84 Plus.

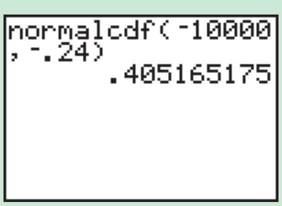
To specify the lower bound in this case, use −10,000.

2nd DISTR

2: normalcdf(

−10000, −.24)

ENTER



Try It Yourself 3

1. Find the cumulative area that corresponds to a z-score of -2.19.

2. Find the cumulative area that corresponds to a z-score of 2.17.

Locate the given z-score and *find the area* that corresponds to it in the Standard Normal Table.

Answer: Page A37

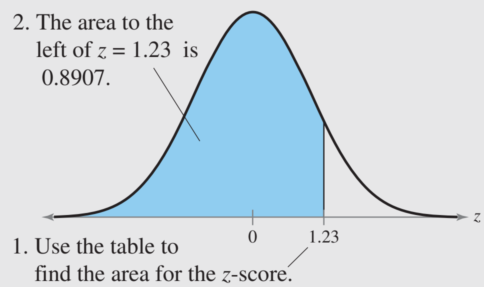
When the z-score is not in the table, use the entry closest to it. If the given z-score is exactly midway between two z-scores, then use the area midway between the corresponding areas.

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

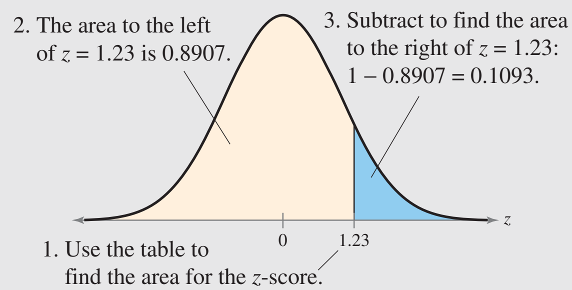
GUIDELINES

Finding Areas Under the Standard Normal Curve

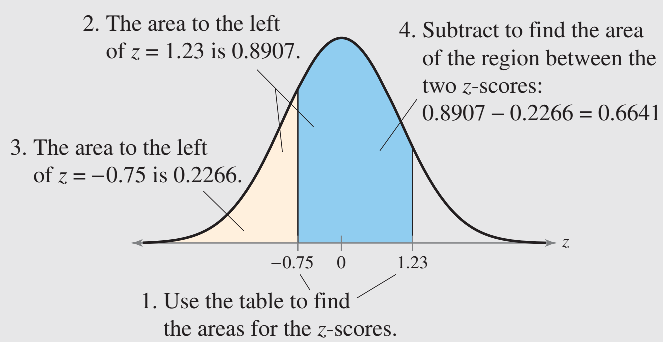
1. Sketch the standard normal curve and shade the appropriate area under the curve.
2. Find the area by following the directions for each case shown.
   a. To find the area to the *left* of z, find the area that corresponds to z in the Standard Normal Table.

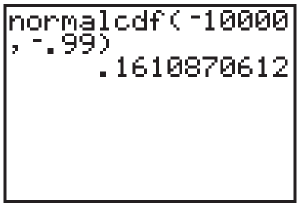


   b. To find the area to the *right* of z, use the Standard Normal Table to find the area that corresponds to z. Then subtract the area from 1.



   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.





Using a TI-83/84 Plus, you can find the area automatically.

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."

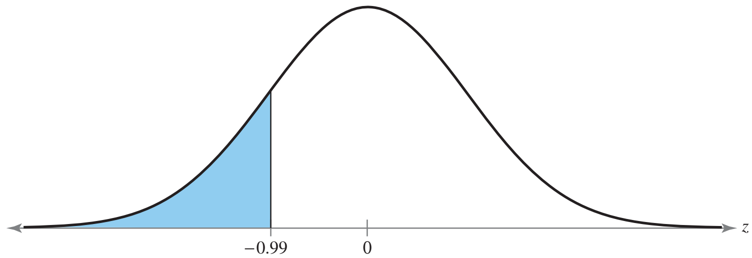
EXAMPLE 4

▸ 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.



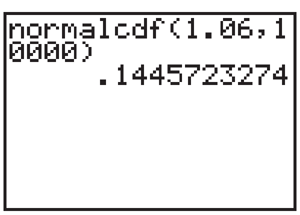
From the Standard Normal Table, this area is equal to 0.1611.

▸ 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 A38



Use 10,000 for the upper bound.

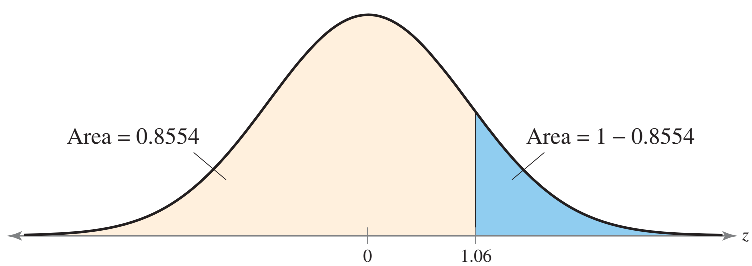
EXAMPLE 5

▸ 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 A38

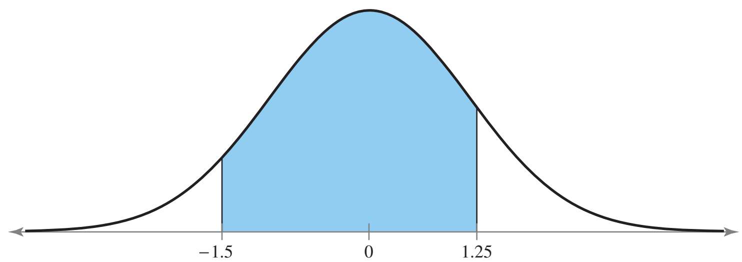
EXAMPLE 6

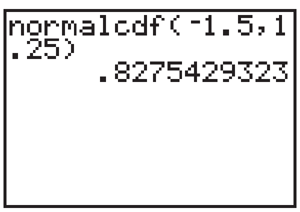
▶ Finding Area Under the Standard Normal Curve

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.





When using technology, your answers may differ slightly from those found using the Standard Normal Table.

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 $z = -2.165$ 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 $z = -2.165$.
c. *Subtract* the smaller area from the larger area.
d. *Interpret* the results.

Answer: Page A38

Recall that in Section 2.4 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.

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. In a normal distribution, which is greater, the mean or the median? Explain.

3. What is the total area under the normal curve?

4. What do the inflection points on a normal distribution represent? Where do they occur?

5. Draw two normal curves that have the same mean but different standard deviations. Describe the similarities and differences.

6. Draw two normal curves that have different means but the same standard deviation. Describe the similarities and differences.

7. What is the mean of the standard normal distribution? What is the standard deviation of the standard normal distribution?

8. Describe how you can transform a nonstandard normal distribution to a standard normal distribution.

9. **Getting at the Concept** Why is it correct to say "a" normal distribution and "the" standard normal distribution?

10. **Getting at the Concept** If a z-score is 0, which of the following must be true? Explain your reasoning.

(a) The mean is 0.

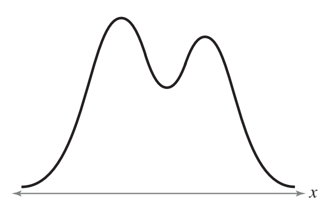
(b) The corresponding x-value is 0.

(c) The corresponding x-value is equal to the mean.

Graphical Analysis *In Exercises 11–16, determine whether the graph could represent a variable with a normal distribution. Explain your reasoning.*

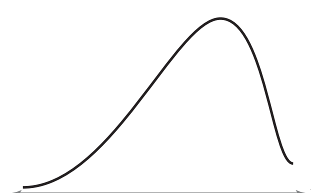
11.

12.

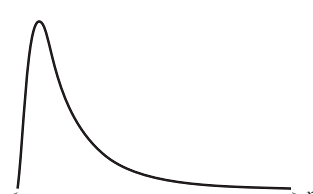


13.

14.



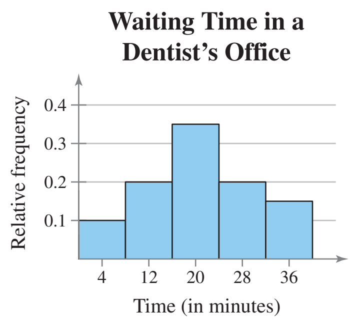
15.



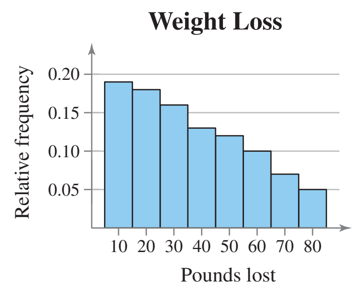
16.

Graphical Analysis *In Exercises 17 and 18, determine whether the histogram represents data with a normal distribution. Explain your reasoning.*

17.



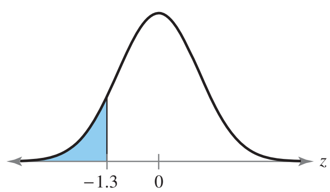
18.



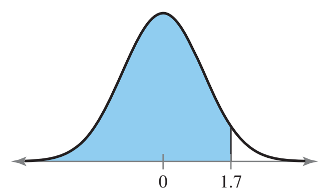
USING AND INTERPRETING CONCEPTS

Graphical Analysis *In Exercises 19–24, find the area of the indicated region under the standard normal curve. If convenient, use technology to find the area.*

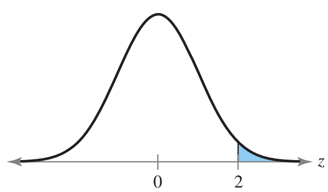
19.



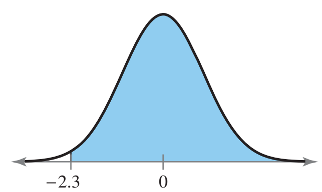
20.



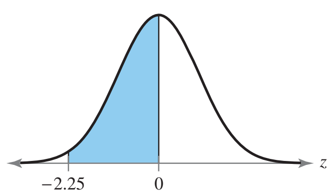
21.



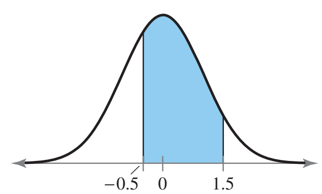
22.



23.



24.



Finding Area *In Exercises 25–38, find the indicated area under the standard normal curve. If convenient, use technology to find the area.*

25. To the left of $z = 0.08$

26. To the right of $z = -3.16$

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

28. To the left of $z = 1.365$

29. To the right of $z = -0.65$

30. To the right of $z = 3.25$

31. To the right of $z = -0.355$

32. To the right of $z = 1.615$

33. Between $z = 0$ and $z = 2.86$

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

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

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

37. To the left of $z = -1.28$ and to the right of $z = 1.28$

38. To the left of $z = -1.96$ and to the right of $z = 1.96$

39. Manufacturer Claims You work for a consumer watchdog publication and are testing the advertising claims of a tire manufacturer. The manufacturer claims that the life spans of the tires are normally distributed, with a mean of 40,000 miles and a standard deviation of 4000 miles. You test 16 tires and get the following life spans.

48,778	41,046	29,083	36,394	32,302	42,787	41,972	37,229
25,314	31,920	38,030	38,445	30,750	38,886	36,770	46,049

(a) Draw a frequency histogram to display these data. Use five classes. Is it reasonable to assume that the life spans 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 manufacturer's claim. Discuss the differences.

40. Milk Consumption You are performing a study about weekly per capita milk consumption. A previous study found weekly per capita milk consumption to be normally distributed, with a mean of 48.7 fluid ounces and a standard deviation of 8.6 fluid ounces. You randomly sample 30 people and find their weekly milk consumptions to be as follows.

40	45	54	41	43	31	47	30	33	37	48	57	52	45	38
65	25	39	53	51	58	52	40	46	44	48	61	47	49	57

(a) Draw a frequency histogram to display these data. Use seven classes. Is it reasonable to assume that the consumptions are normally distributed? Why?

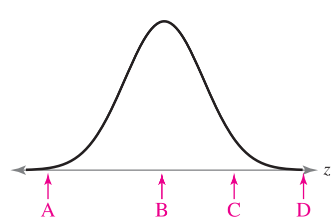
(b) Find the mean and standard deviation of your sample.

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

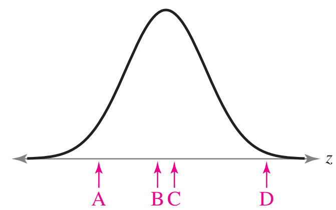
Computing and Interpreting *z*-Scores of Normal Distributions

In Exercises 41–44, 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 the values 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.

41. Blood Pressure The systolic blood pressures of a sample of adults are normally distributed, with a mean pressure of 115 millimeters of mercury and a standard deviation of 3.6 millimeters of mercury. The systolic blood pressures of four adults selected at random are 121 millimeters of mercury, 113 millimeters of mercury, 105 millimeters of mercury, and 127 millimeters of mercury.



42. Cereal Boxes The weights of the contents of cereal boxes are normally distributed, with a mean weight of 12 ounces and a standard deviation of 0.05 ounce. The weights of the contents of four cereal boxes selected at random are 12.01 ounces, 11.92 ounces, 12.12 ounces, and 11.99 ounces.



43. SAT 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 1509 and the standard deviation was 312. The test scores of four students selected at random are 1924, 1241, 2202, and 1392. *(Source: The College Board)*

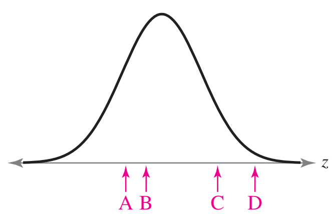


FIGURE FOR EXERCISE 43

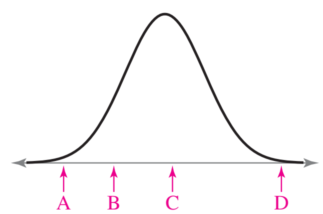
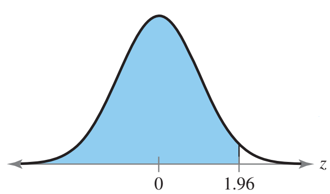


FIGURE FOR EXERCISE 44

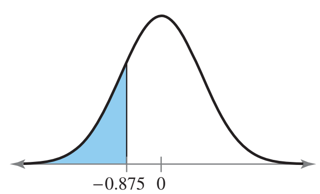
44. ACT Scores The ACT is an exam used by colleges and universities to evaluate undergraduate applicants. The test scores are normally distributed. In a recent year, the mean test score was 21.1 and the standard deviation was 5.0. The test scores of four students selected at random are 15, 22, 9, and 35. *(Source: ACT, Inc.)*

Graphical Analysis *In Exercises 45–50, find the probability of z occurring in the indicated region. If convenient, use technology to find the probability.*

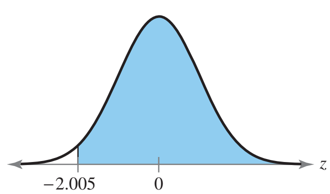
45.



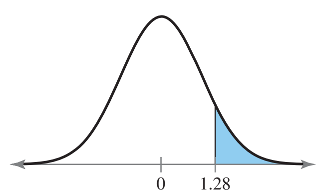
46.



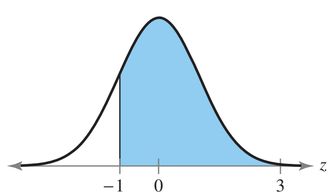
47.



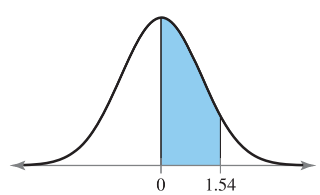
48.



49.



50.



Finding Probabilities *In Exercises 51–60, find the indicated probability using the standard normal distribution. If convenient, use technology to find the probability.*

51. $P(z < 1.45)$ **52.** $P(z < -0.18)$ **53.** $P(z > 2.175)$

54. $P(z > -1.85)$ **55.** $P(-0.89 < z < 0)$ **56.** $P(0 < z < 0.525)$

57. $P(-1.65 < z < 1.65)$ **58.** $P(-1.54 < z < 1.54)$

59. $P(z < -2.58 \text{ or } z > 2.58)$ **60.** $P(z < -1.54 \text{ or } z > 1.54)$

EXTENDING CONCEPTS

61. 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.

62. 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.

63. Uniform Distribution Another continuous distribution is the **uniform distribution.** An example is $f(x) = 1$ for $0 \le x \le 1$. The mean of the distribution for this example is 0.5 and the standard deviation is approximately 0.29. The graph of the 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

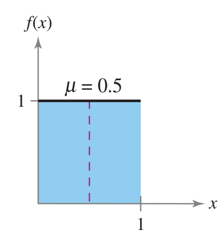
$$f(x) = \frac{1}{b - a}.$$

The mean is

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

and the standard deviation is

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



(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.

64. 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.

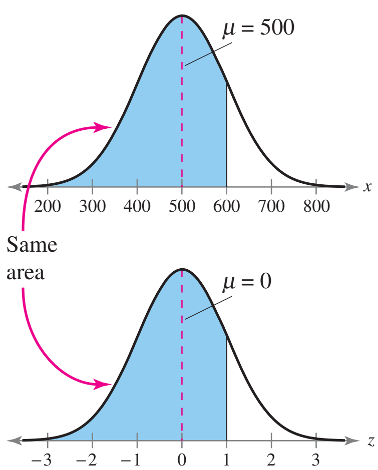
(b) Find the probability that x falls between 12 and 15.

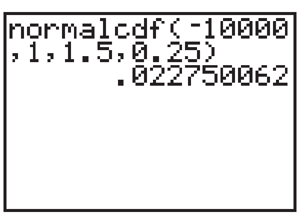
(c) Find the probability that x falls between 13 and 18.

5.2 Normal Distributions: Finding Probabilities

WHAT YOU SHOULD LEARN

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





In Example 1, you can use a TI-83/84 Plus to find the probability automatically.

STUDY TIP

Another way to write the answer to Example 1 is $P(x < 1) = 0.0228$.

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, you can 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 SC Report 20

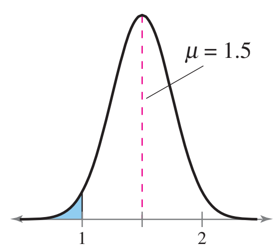
Finding Probabilities for Normal Distributions

A survey indicates that people use their cellular phones an average of 1.5 years before buying a new one. The standard deviation is 0.25 year. A cellular phone user is selected at random. Find the probability that the user will use their current phone for less than 1 year before buying a new one. Assume that the variable x is normally distributed. *(Adapted from Fonebak)*

Solution

The graph shows a normal curve with $\mu = 1.5$ and $\sigma = 0.25$ and a shaded area for x less than 1. The z-score that corresponds to 1 year is

$$z = \frac{x - \mu}{\sigma} = \frac{1 - 1.5}{0.25} = -2.$$



Age of cellular phone (in years)

The Standard Normal Table shows that $P(z < -2) = 0.0228$. The probability that the user will use their cellular phone for less than 1 year before buying a new one is 0.0228.

Interpretation So, 2.28% of cellular phone users will use their cellular phone for less than 1 year before buying a new one. Because 2.28% is less than 5%, this is an unusual event.

Try It Yourself 1

The average speed of vehicles traveling on a stretch of highway is 67 miles per hour with a standard deviation of 3.5 miles per hour. A vehicle is selected at random. What is the probability that it is violating the 70 mile per hour speed limit? Assume the speeds are normally distributed.

a. *Sketch* a graph.
b. *Find the z-score* that corresponds to 70 miles per hour.
c. *Find the area* to the right of that z-score.
d. *Interpret* the results.

Answer: Page A38

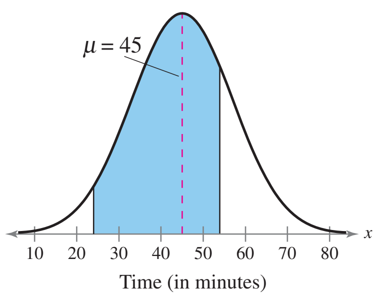
EXAMPLE 2

Finding Probabilities for Normal Distributions

A survey indicates that for each trip to the supermarket, a shopper spends an average of 45 minutes with a standard deviation of 12 minutes in the store. The lengths of time spent in the store are normally distributed and are 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) Interpret your answer 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



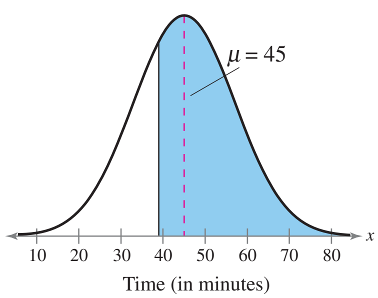
1. (a) The graph at the left shows a normal curve with $\mu = 45$ minutes and $\sigma = 12$ minutes. The area for x between 24 and 54 minutes is shaded. The z-scores that correspond to 24 minutes and to 54 minutes are

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

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) ***Interpretation*** 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) ***Interpretation*** 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 in Example 2 will be in the supermarket between 33 and 60 minutes?

a. *Sketch* a graph.
b. *Find the z-scores* that correspond to 33 minutes and 60 minutes.
c. *Find the cumulative area* for each z-score and *subtract* the smaller area from the larger area.
d. *Interpret* your answer if 150 shoppers enter the store. How many shoppers would you expect to be in the store between 33 and 60 minutes?

Answer: Page A38

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-83/84 Plus.

EXAMPLE 3

Using Technology to Find Normal Probabilities

Triglycerides are a type of fat in the bloodstream. The mean triglyceride level in the United States is 134 milligrams per deciliter. Assume the triglyceride levels of the population of the United States are normally distributed, with a standard deviation of 35 milligrams per deciliter. You randomly select a person from the United States. What is the probability that the person's triglyceride level is less than 80? Use a technology tool to find the probability. *(Adapted from University of Maryland Medical Center)*

Solution

MINITAB, Excel, and the TI-83/84 Plus 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 = 134 and standard deviation = 35

x	P(X <= x)
80	0.0614327

EXCEL

	A	B	C
1	NORMDIST(80,134,35,TRUE)		
2			0.06143272

TI-83/84 PLUS

normalcdf(-10000,80,134,35)
.0614327356

From the displays, you can see that the probability that the person's triglyceride level is less than 80 is about 0.0614, or 6.14%.

Try It Yourself 3

A person from the United States is selected at random. What is the probability that the person's triglyceride level is between 100 and 150? 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.

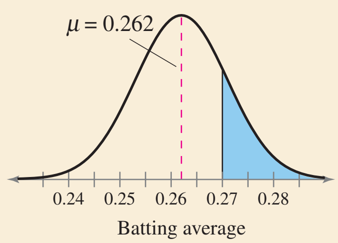
Answer: Page A38

Example 3 shows only one of several ways to find normal probabilities using MINITAB, Excel, and the TI-83/84 Plus.

PICTURING THE WORLD

In baseball, a batting average is the number of hits divided by the number of at-bats. The batting averages of all 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 0.262 and the standard deviation is 0.009. (Adapted from ESPN)

Major League Baseball



What percent of the players have a batting average of 0.270 or greater? If there are 40 players on a roster, how many would you expect to have a batting average of 0.270 or greater?

5.2 EXERCISES

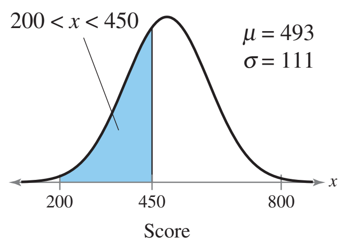
BUILDING BASIC SKILLS AND VOCABULARY

Computing Probabilities *In Exercises 1–6, assume the random variable x is normally distributed with mean* $\mu = 174$ *and standard deviation* $\sigma = 20$. *Find the indicated probability.*

1. $P(x < 170)$
2. $P(x < 200)$
3. $P(x > 182)$
4. $P(x > 155)$
5. $P(160 < x < 170)$
6. $P(172 < x < 192)$

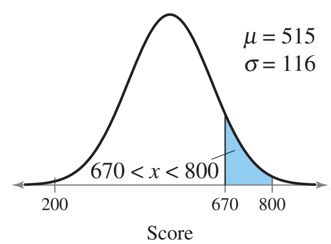
Graphical Analysis *In Exercises 7–12, assume a member is selected at random from the population represented by the graph. Find the probability that the member selected at random is from the shaded area of the graph. Assume the variable x is normally distributed.*

7. **SAT Writing Scores**



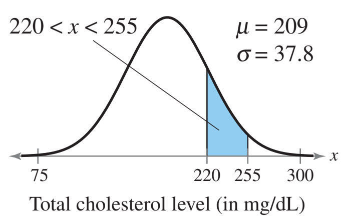
(Source: The College Board)

8. **SAT Math Scores**



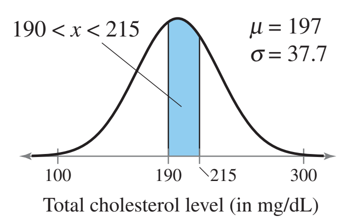
(Source: The College Board)

9. **U.S. Men Ages 35–44: Total Cholesterol**



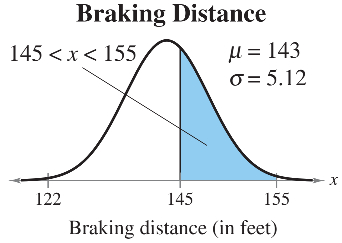
(Adapted from National Center for Health Statistics)

10. **U.S. Women Ages 35–44: Total Cholesterol**



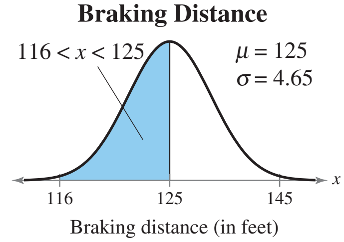
(Adapted from National Center for Health Statistics)

11. **Ford Fusion: Braking Distance**



(Adapted from Consumer Reports)

12. **Hyundai Elantra: Braking Distance**



(Adapted from Consumer Reports)

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 heights 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 69.9 inches and a standard deviation of 3.0 inches. A study participant is randomly selected. *(Adapted from 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.
(d) Can any of these events be considered unusual? Explain your reasoning.

14. Heights of Women A survey was conducted to measure the heights of U.S. women. In the survey, respondents were grouped by age. In the 20–29 age group, the heights were normally distributed, with a mean of 64.3 inches and a standard deviation of 2.6 inches. A study participant is randomly selected. *(Adapted from U.S. National Center for Health Statistics)*

(a) Find the probability that her height is less than 56.5 inches.
(b) Find the probability that her height is between 61 and 67 inches.
(c) Find the probability that her height is more than 70.5 inches.
(d) Can any of these events be considered unusual? Explain your reasoning.

15. ACT English Scores In a recent year, the ACT scores for the English portion of the test were normally distributed, with a mean of 20.6 and a standard deviation of 6.3. A high school student who took the English portion of the ACT is randomly selected. *(Source: ACT, Inc.)*

(a) Find the probability that the student's ACT score is less than 15.
(b) Find the probability that the student's ACT score is between 18 and 25.
(c) Find the probability that the student's ACT score is more than 34.
(d) Can any of these events be considered unusual? Explain your reasoning.

16. Beagles The weights of adult male beagles are normally distributed, with a mean of 25 pounds and a standard deviation of 3 pounds. A beagle is randomly selected.

(a) Find the probability that the beagle's weight is less than 23 pounds.
(b) Find the probability that the weight is between 24.5 and 25 pounds.
(c) Find the probability that the beagle's weight is more than 30 pounds.
(d) Can any of these events be considered unusual? Explain your reasoning.

17. Computer Usage A survey was conducted to measure the number of hours per week adults in the United States spend on their computers. In the survey, the numbers 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 number of hours spent on the computer by the participant is less than 5 hours per week.
(b) Find the probability that the number of hours spent on the computer by the participant is between 5.5 and 9.5 hours per week.
(c) Find the probability that the number of hours spent on the computer by the participant is more than 10 hours per week.

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 \$70.

(b) Find the probability that the utility bill is between \$90 and \$120.

(c) Find the probability that the utility bill is more than \$140.

19. Computer Lab Schedule The times per week a student uses a lab computer are 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 5 and 7 hours per week.

(c) Find the probability that the student uses a lab computer more than 8 hours per week.

20. Health Club Schedule The times per workout an athlete uses a stairclimber are 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 20 and 28 minutes.

(c) Find the probability that the athlete uses a stairclimber for more than 30 minutes.

Using Normal Distributions *In Exercises 21–28, answer the questions about the specified normal distribution.*

21. SAT Writing Scores Use the normal distribution of SAT writing scores in Exercise 7 for which the mean is 493 and the standard deviation is 111.

(a) What percent of the SAT writing scores are less than 600?

(b) If 1000 SAT writing 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 515 and the standard deviation is 116.

(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 men's total cholesterol levels in Exercise 9 for which the mean is 209 milligrams per deciliter and the standard deviation is 37.8 milligrams per deciliter.

(a) What percent of the men have a total cholesterol level less than 225 milligrams per deciliter of blood?

(b) If 250 U.S. men in the 35–44 age group are randomly selected, about how many would you expect to have a total cholesterol level greater than 260 milligrams per deciliter of blood?

24. Cholesterol Use the normal distribution of women's total cholesterol levels in Exercise 10 for which the mean is 197 milligrams per deciliter and the standard deviation is 37.7 milligrams per deciliter.

(a) What percent of the women have a total cholesterol level less than 217 milligrams per deciliter of blood?

(b) If 200 U.S. women in the 35–44 age group are randomly selected, about how many would you expect to have a total cholesterol level greater than 185 milligrams per deciliter of blood?

25. 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 their 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 their computer?

26. 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?

27. Battery Life Spans The life spans of batteries are 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.

28. Peanuts Assume the mean annual consumptions of peanuts are 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.

SC *In Exercises 29 and 30, use the StatCrunch normal calculator to find the indicated probabilities.*

29. Soft Drink Machine The amounts a soft drink machine is designed to dispense for each drink are normally distributed, with a mean of 12 fluid ounces and a standard deviation of 0.2 fluid ounce. A drink is randomly selected.

(a) Find the probability that the drink is less than 11.9 fluid ounces.

(b) Find the probability that the drink is between 11.8 and 11.9 fluid ounces.

(c) Find the probability that the drink is more than 12.3 fluid ounces. Can this be considered an unusual event? Explain your reasoning.

30. Machine Parts The thicknesses of washers produced by a machine are normally distributed, with a mean of 0.425 inch and a standard deviation of 0.005 inch. A washer is randomly selected.

(a) Find the probability that the washer is less than 0.42 inch thick.

(b) Find the probability that the washer is between 0.40 and 0.42 inch thick.

(c) Find the probability that the washer is more than 0.44 inch thick. Can this be considered an unusual event? Explain your reasoning.

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.*

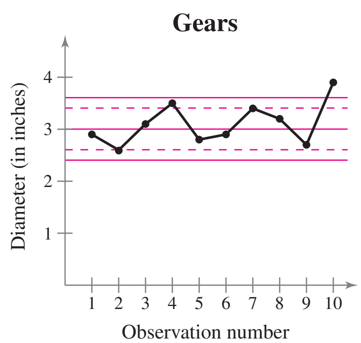
(1) A point lies beyond three standard deviations of the mean.

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

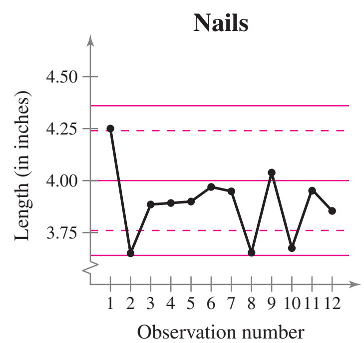
(3) At least two of three consecutive points lie more than two standard deviations from the mean.

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.

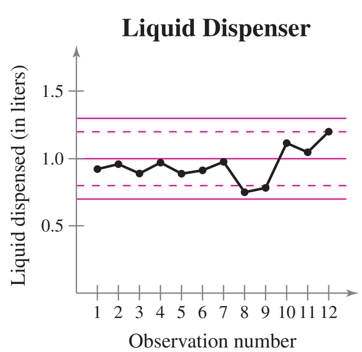
31. A gear has been designed to have a diameter of 3 inches. The standard deviation of the process is 0.2 inch.



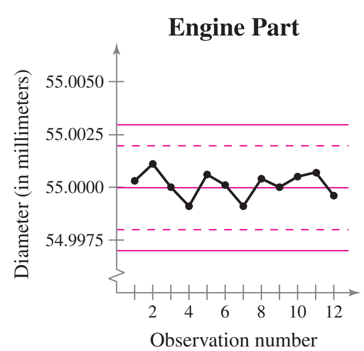
32. A nail has been designed to have a length of 4 inches. The standard deviation of the process is 0.12 inch.



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.

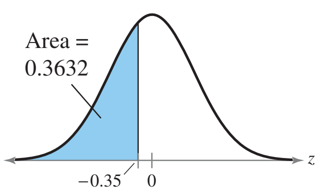
But what if you are given a probability and want to find a value? For instance, a university might want to know the lowest test score a student can have on an entrance exam and still be in the top 10%, or a medical researcher might want to know the cutoff values for selecting the middle 90% of patients 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.

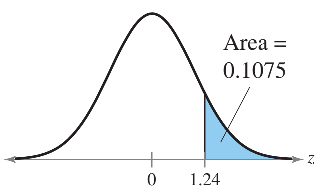
EXAMPLE 1

▸ 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

You can also use a computer or calculator to find the z-scores that correspond to the given cumulative areas, as shown in the margin.

STUDY TIP

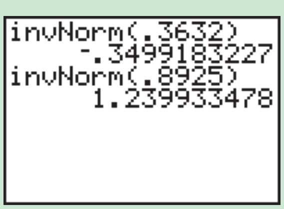
Here are instructions for finding the z-score that corresponds to a given area on a TI-83/84 Plus.

2nd DISTR

3: invNorm(

Enter the cumulative area.

ENTER



▸ 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 A38*

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.

In Section 2.5, you learned that percentiles divide a data set into 100 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

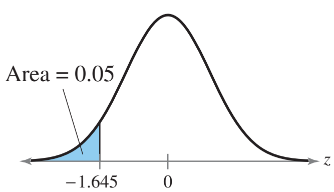
▸ Finding a *z*-Score Given a Percentile

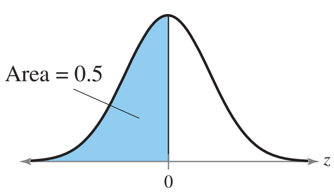
Find the z-score that corresponds to each percentile.

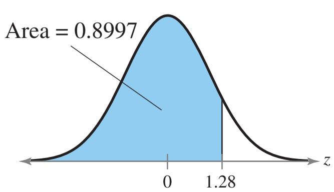
1. P_5

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 upper 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 middle 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.

3. To find the z-score that corresponds to P_{90}, find the z-score that corresponds to an area of 0.9 (see lower 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 about 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 A38*

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

A veterinarian records the weights of cats treated at a clinic. The weights are normally distributed, with a mean of 9 pounds and a standard deviation of 2 pounds. Find the weights x corresponding to z-scores of 1.96, -0.44, and 0. Interpret your results.

Solution

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

$z = 1.96$: $\quad x = 9 + 1.96(2) = 12.92$ pounds

$z = -0.44$: $\quad x = 9 + (-0.44)(2) = 8.12$ pounds

$z = 0$: $\quad x = 9 + 1.96(0) = 9$ pounds

Interpretation You can see that 12.92 pounds is above the mean, 8.12 pounds is below the mean, and 9 pounds is equal to the mean.

Try It Yourself 3

A veterinarian records the weights of dogs treated at a clinic. The weights are normally distributed, with a mean of 52 pounds and a standard deviation of 15 pounds. Find the weights x corresponding to z-scores of -2.33, 3.10, and 0.58. Interpret your results.

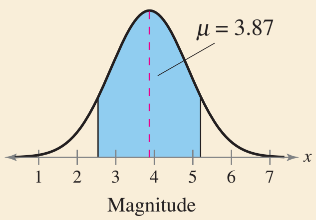
a. *Identify* μ and σ of the normal distribution.
b. *Transform* each z-score to an x-value.
c. *Interpret* the results.

Answer: Page A38

PICTURING THE WORLD

According to the United States Geological Survey, the mean magnitude of worldwide earthquakes in a recent year was about 3.87. The magnitude of worldwide earthquakes can be approximated by a normal distribution. Assume the standard deviation is 0.81. (Adapted from United States Geological Survey)

Worldwide Earthquakes in 2008



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

STUDY TIP

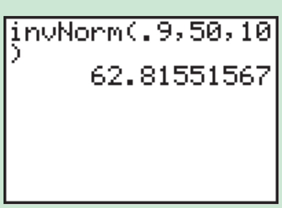
Here are instructions for finding a specific x-value for a given probability on a TI-83/84 Plus.

2nd DISTR

3: invNorm(

Enter the values for the area under the normal distribution, the specified mean, and the specified standard deviation separated by commas.

ENTER



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 Examples 4 and 5.

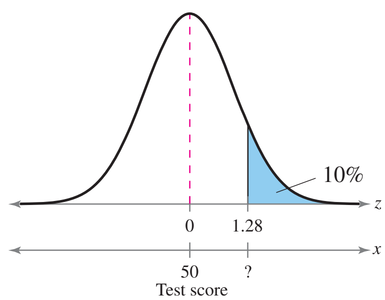
EXAMPLE 4 SC Report 21

Finding a Specific Data Value

Scores for the California Peace Officer Standards and Training test are normally distributed, with a mean of 50 and a standard deviation of 10. An agency will only hire applicants with scores in the top 10%. What is the lowest score you can earn and still be eligible to be hired by the agency? *(Source: State of California)*

Solution

Exam scores in the top 10% correspond to the shaded region shown.



A test score in the top 10% is any score above the 90th percentile. To find the score that represents the 90th percentile, you must first find the z-score that corresponds to a cumulative area of 0.9. From the Standard Normal Table, you can find that the area closest to 0.9 is 0.8997. So, the z-score that corresponds to an area of 0.9 is $z = 1.28$. Using the equation $x = \mu + z\sigma$, you have

$$\begin{aligned} x &= \mu + z\sigma \\ &= 50 + 1.28(10) \\ &\approx 62.8. \end{aligned}$$

Interpretation The lowest score you can earn and still be eligible to be hired by the agency is about 63.

Try It Yourself 4

The braking distances of a sample of Nissan Altimas are normally distributed, with a mean of 129 feet and a standard deviation of 5.18 feet. What is the longest braking distance one of these Nissan Altimas could have and still be in the bottom 1%? *(Adapted from Consumer Reports)*

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 A38

EXAMPLE 5

SC Report 22

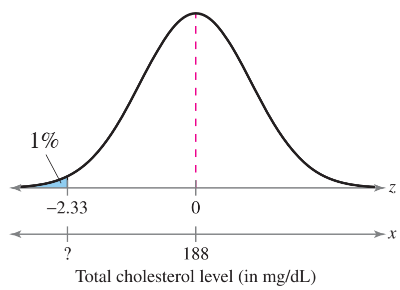
Finding a Specific Data Value

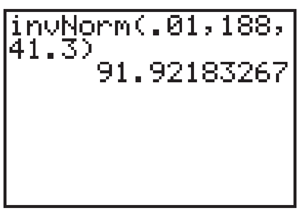
In a randomly selected sample of women ages 20–34, the mean total cholesterol level is 188 milligrams per deciliter with a standard deviation of 41.3 milligrams per deciliter. Assume the total cholesterol levels are normally distributed. Find the highest total cholesterol level a woman in this 20–34 age group can have and still be in the bottom 1%. *(Adapted from National Center for Health Statistics)*

Solution

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

Total Cholesterol Levels in Women Ages 20–34





Using a TI-83/84 Plus, you can find the highest total cholesterol level automatically.

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

$$\begin{aligned} x &= \mu + z\sigma \\ &= 188 + (-2.33)(41.3) \\ &\approx 91.77. \end{aligned}$$

Interpretation The value that separates the lowest 1% of total cholesterol levels for women in the 20–34 age group from the highest 99% is about 92 milligrams per deciliter.

Try It Yourself 5

The lengths of time employees have worked at a corporation are 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 A38

5.3 EXERCISES

BUILDING BASIC SKILLS AND VOCABULARY

In Exercises 1–16, 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. If convenient, use technology to find the z-score.

1. 0.2090 **2.** 0.4364 **3.** 0.9916 **4.** 0.7995

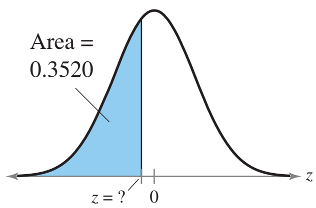
5. 0.05 **6.** 0.85 **7.** 0.94 **8.** 0.0046

9. P_{15} **10.** P_{30} **11.** P_{88} **12.** P_{67}

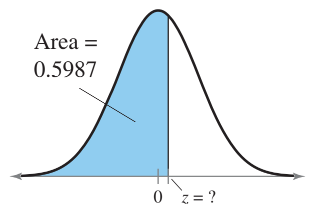
13. P_{25} **14.** P_{40} **15.** P_{75} **16.** P_{80}

Graphical Analysis *In Exercises 17–22, find the indicated z-score(s) shown in the graph. If convenient, use technology to find the z-score(s).*

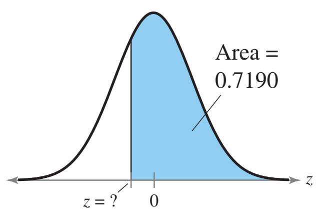
17.



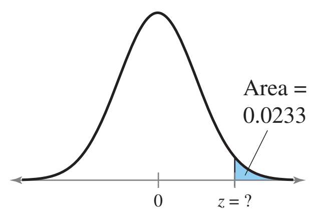
18.



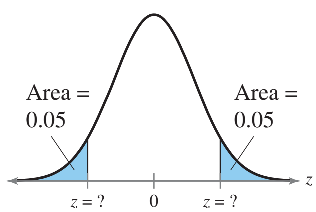
19.



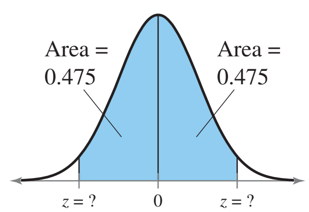
20.



21.



22.



In Exercises 23–30, find the indicated z-score.

23. Find the z-score that has 11.9% of the distribution's area to its left.

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

25. Find the z-score that has 11.9% of the distribution's area to its right.

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

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

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

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

30. 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 31–36, answer the questions about the specified normal distribution.*

31. Heights of Women In a survey of women in the United States (ages 20–29), the mean height was 64.3 inches with a standard deviation of 2.6 inches. *(Adapted from National Center for Health Statistics)*

(a) What height represents the 95th percentile?

(b) What height represents the first quartile?

32. Heights of Men In a survey of men in the United States (ages 20–29), the mean height was 69.9 inches with a standard deviation of 3.0 inches. *(Adapted from National Center for Health Statistics)*

(a) What height represents the 90th percentile?

(b) What height represents the first quartile?

33. Heart Transplant Waiting Times The time spent (in days) waiting for a heart transplant for people ages 35–49 in a recent year can be approximated by a normal distribution, as shown in the graph. *(Adapted from Organ Procurement and Transplantation Network)*

(a) What waiting time represents the 5th percentile?

(b) What waiting time represents the third quartile?

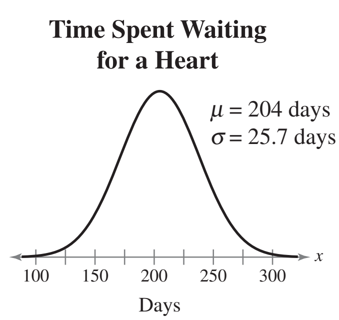


FIGURE FOR EXERCISE 33

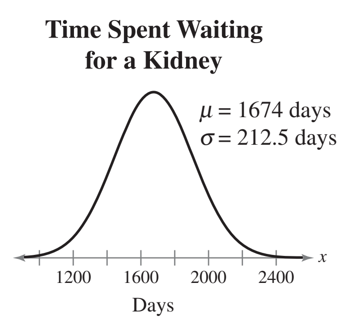


FIGURE FOR EXERCISE 34

34. Kidney Transplant Waiting Times The time spent (in days) waiting for a kidney transplant for people ages 35–49 in a recent year can be approximated by a normal distribution, as shown in the graph. *(Adapted from Organ Procurement and Transplantation Network)*

(a) What waiting time represents the 80th percentile?

(b) What waiting time represents the first quartile?

35. Sleeping Times of Medical Residents The average time spent sleeping (in hours) for a group of medical residents at a hospital can be approximated by a normal distribution, as shown in the graph. *(Source: National Institute of Occupational Safety and Health, Japan)*

(a) What is the shortest time spent sleeping that would still place a resident in the top 5% of sleeping times?

(b) Between what two values does the middle 50% of the sleep times lie?

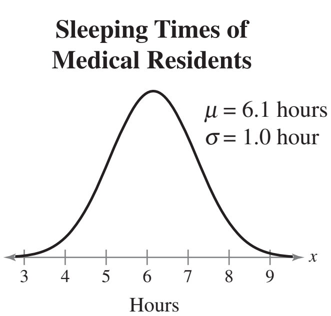


FIGURE FOR EXERCISE 35

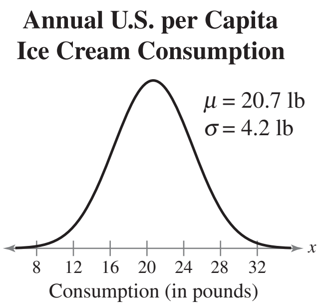


FIGURE FOR EXERCISE 36

36. 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 largest annual per capita consumption of ice cream that can be in the bottom 10% of consumptions?

(b) Between what two values does the middle 80% of the consumptions lie?

37. Bags of Baby Carrots The weights of bags of baby carrots are normally distributed, with a mean of 32 ounces and a standard deviation of 0.36 ounce. Bags in the upper 4.5% are too heavy and must be repackaged. What is the most a bag of baby carrots can weigh and not need to be repackaged?

38. Vending Machine A vending machine dispenses coffee into an eight-ounce cup. The amounts of coffee dispensed into the cup are 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?

SC *In Exercises 39 and 40, use the StatCrunch normal calculator to find the indicated values.*

39. Apples The annual per capita consumption of fresh apples (in pounds) in the United States can be approximated by a normal distribution, with a mean of 16.2 pounds and a standard deviation of 4 pounds. *(Adapted from U.S. Department of Agriculture)*

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

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

40. Oranges The annual per capita consumption of fresh oranges (in pounds) in the United States can be approximated by a normal distribution, with a mean of 9.9 pounds and a standard deviation of 2.5 pounds. *(Adapted from U.S. Department of Agriculture)*

(a) What is the smallest annual per capita consumption of oranges that can be in the top 10% of consumptions?

(b) What is the largest annual per capita consumption of oranges that can be in the bottom 5% of consumptions?

EXTENDING CONCEPTS

41. 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?

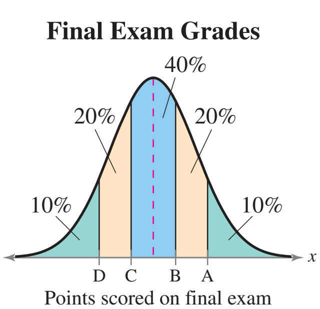


FIGURE FOR EXERCISE 42

42. 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 A's, the next 20% receive B's, the middle 40% receive C's, the next 20% receive D's, and the bottom 10% receive F's. Find the lowest score on the final exam that would qualify a student for an A, a B, a C, and a D.

CASE STUDY

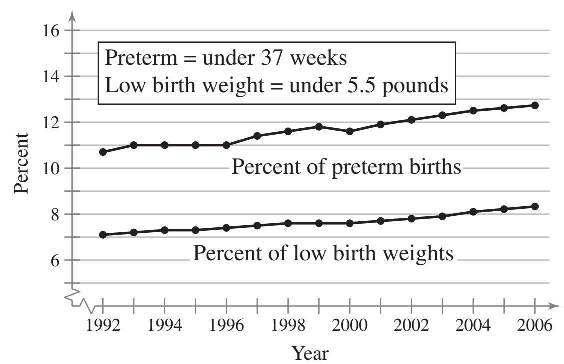
Birth Weights in America

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 in the table below.

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 below, the problem of low birth weights increased from 1992 to 2006.

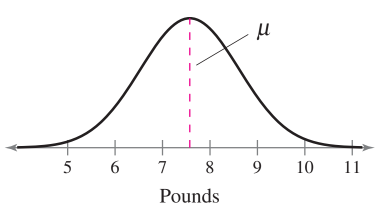
Gestation period	Mean birth weight	Standard deviation
Under 28 weeks	1.90 lb	1.22 lb
28 to 31 weeks	4.12 lb	1.87 lb
32 to 33 weeks	5.14 lb	1.57 lb
34 to 36 weeks	6.19 lb	1.29 lb
37 to 39 weeks	7.29 lb	1.08 lb
40 weeks	7.66 lb	1.04 lb
41 weeks	7.75 lb	1.07 lb
42 weeks and over	7.57 lb	1.11 lb



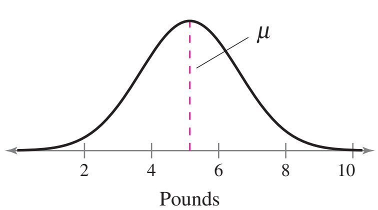
EXERCISES

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

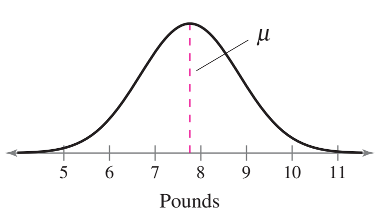
(a)



(b)



(c)



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

(a) Under 28 weeks (b) 32 to 33 weeks
(c) 40 weeks (d) 42 weeks and over

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

(a) Under 28 weeks (b) 34 to 36 weeks
(c) 41 weeks (d) 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) Under 28 weeks (b) 28 to 31 weeks
(c) 34 to 36 weeks (d) 37 to 39 weeks

5. A birth weight of less than 3.25 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) 28 to 31 weeks
(c) 32 to 33 weeks (d) 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

▸ 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.

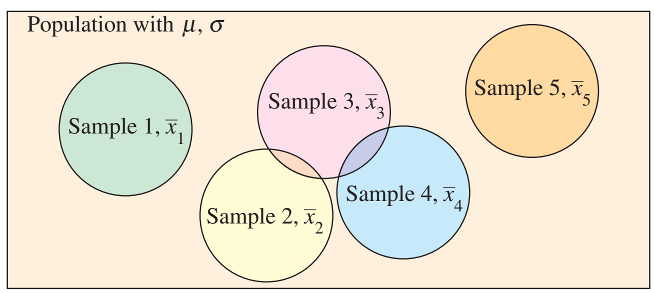
DEFINITION

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.** Every sample statistic has a sampling distribution.

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*.

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 $\overline{x}_1$; the mean of Sample 2 is $\overline{x}_2$; and so on. The sampling distribution of the sample means for samples of size n for this population consists of $\overline{x}_1$, $\overline{x}_2$, $\overline{x}_3$, and so on. If the samples are drawn with replacement, an infinite number of samples can be drawn from the population.



PROPERTIES OF SAMPLING DISTRIBUTIONS OF SAMPLE MEANS

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

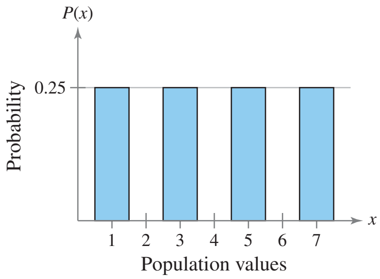
$$\mu_{\overline{x}} = \mu$$

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

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

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

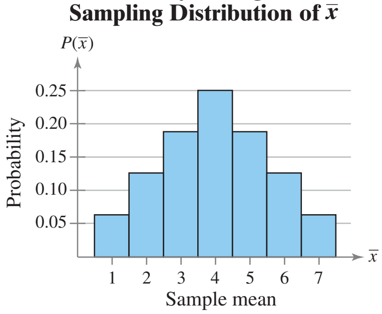
Probability Histogram of Population of x



Probability Distribution of Sample Means

$\bar{x}$	f	Probability
1	1	$1/16 = 0.0625$
2	2	$2/16 = 0.1250$
3	3	$3/16 = 0.1875$
4	4	$4/16 = 0.2500$
5	3	$3/16 = 0.1875$
6	2	$2/16 = 0.1250$
7	1	$1/16 = 0.0625$

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



To explore this topic further, see Activity 5.4 on page 280.

STUDY TIP

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

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.

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

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

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

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

These results satisfy the properties of sampling distributions because

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

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 of the population.

Answer: Page A38

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.

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, then 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.

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

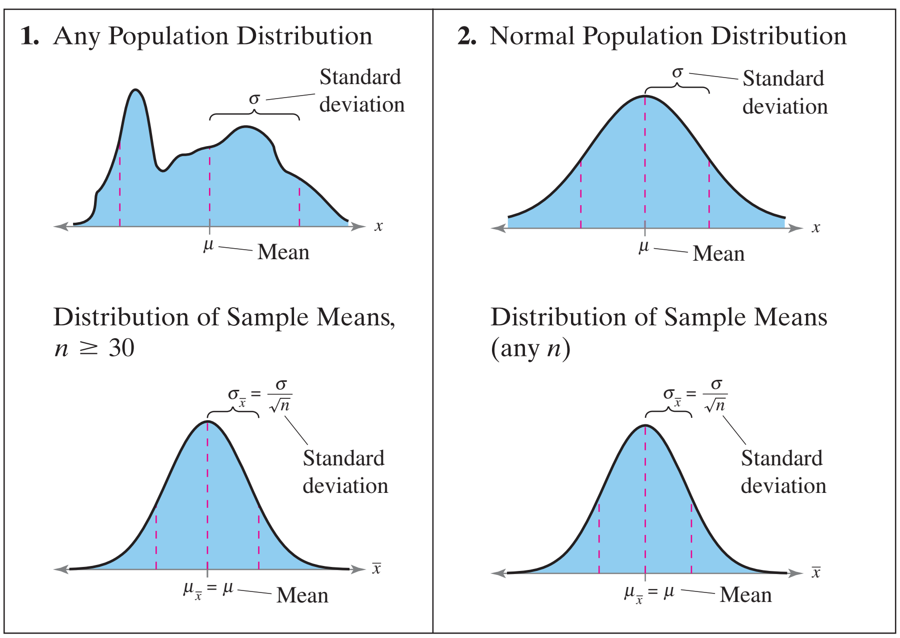
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad \text{Standard deviation}$$

Recall that 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.

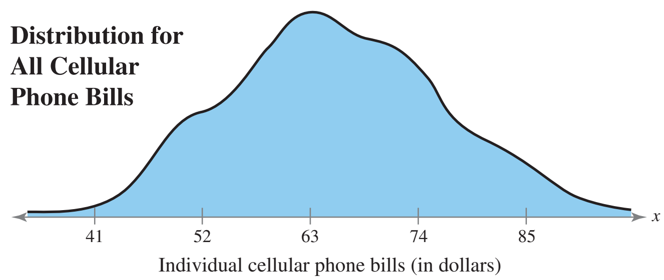
Moreover, the distribution of sample means becomes less and less spread out (tighter concentration about the mean) as the sample size n increases.



EXAMPLE 2

Interpreting the Central Limit Theorem

Cellular phone bills for residents of a city have a mean of \$63 and a standard deviation of \$11, as shown in the following graph. Random samples of 100 cellular 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. *(Adapted from JD Power and Associates)*

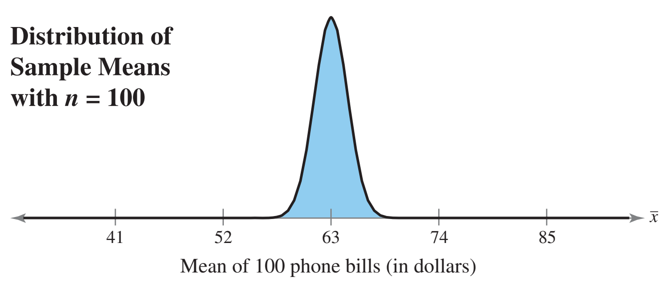


Solution

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,

$$\mu_{\bar{x}} = \mu = 63 \quad \text{and} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{11}{\sqrt{100}} = 1.1.$$

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 $\mu = \$63$ and $\sigma = \$1.10$, as shown in the graph below.



Try It Yourself 2

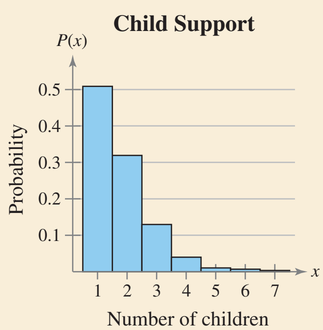
Suppose random samples of size 64 are drawn from the population in Example 2. Find the mean and standard error of the mean of the sampling distribution. Sketch a graph of the sampling distribution and compare it with the sampling distribution in Example 2.

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 A39

PICTURING THE WORLD

In a recent year, there were about 4.8 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 0.8. (Adapted from U.S. Census Bureau)

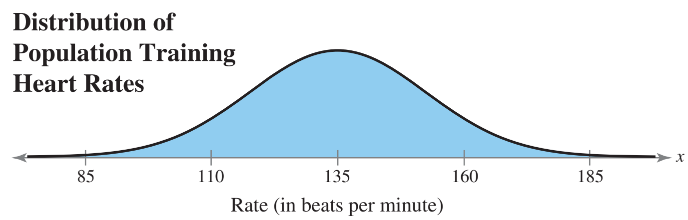


You randomly select 35 parents who receive child support and ask how many children in their custody are receiving child support payments. What is the probability that the mean of the sample is between 1.5 and 1.9 children?

EXAMPLE 3

Interpreting the Central Limit Theorem

Suppose the training heart rates of all 20-year-old athletes are normally distributed, with a mean of 135 beats per minute and standard deviation of 18 beats per minute, 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.

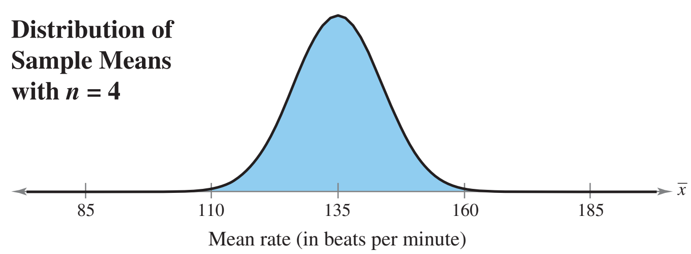


Solution

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,

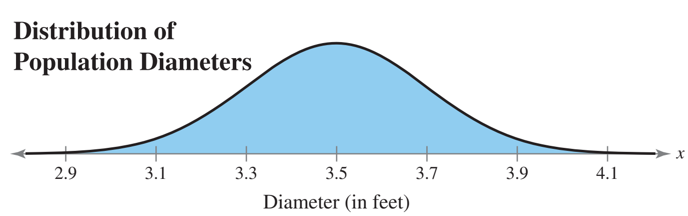
$$\mu_{\bar{x}} = \mu = 135 \text{ beats per minute and } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{18}{\sqrt{4}} = 9 \text{ beats per minute.}$$

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 A39

PROBABILITY AND THE CENTRAL LIMIT THEOREM

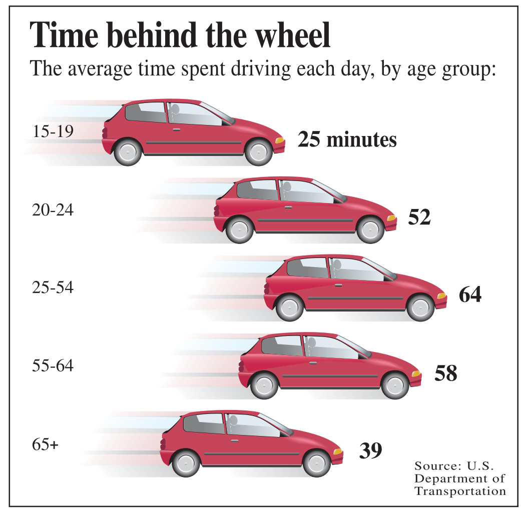
In Section 5.2, you learned how to find the probability that a random variable x will fall in a given interval of population values. In a similar manner, you can find the probability that a sample mean $\bar{x}$ will fall in a given interval of the $\bar{x}$ sampling distribution. To transform $\bar{x}$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

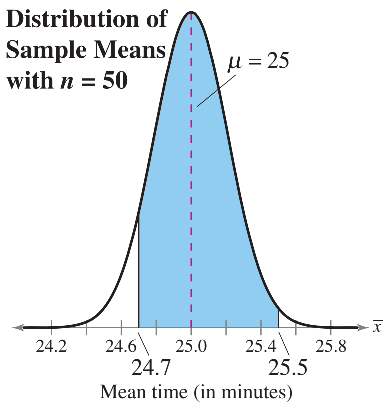
The graph at the right shows the lengths of time people spend driving each day. You randomly select 50 drivers ages 15 to 19. What is the probability that the mean time they spend driving each day is between 24.7 and 25.5 minutes? Assume that $\sigma = 1.5$ minutes.



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

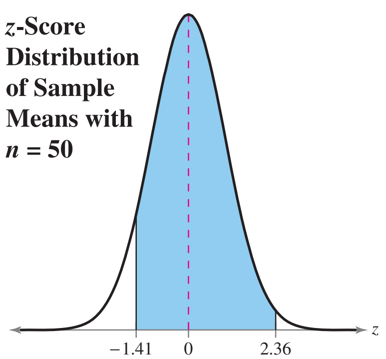
$$\mu_{\bar{x}} = \mu = 25 \text{ minutes} \quad \text{and} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.5}{\sqrt{50}} \approx 0.21213 \text{ minute.}$$



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

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

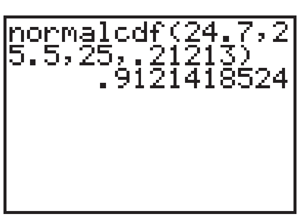
$$z_2 = \frac{25.5 - 25}{1.5/\sqrt{50}} \approx \frac{0.5}{0.21213} \approx 2.36.$$



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.



In Example 4, you can use a TI-83/84 Plus to find the probability automatically once the standard error of the mean is calculated.

STUDY TIP

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

▸ Try It Yourself 4

You randomly select 100 drivers ages 15 to 19 from Example 4. 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.

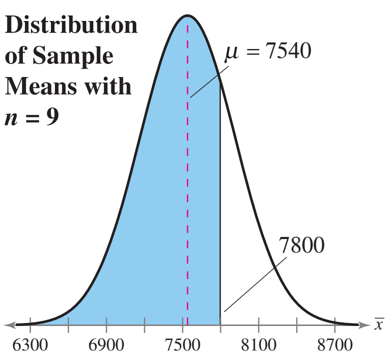
a. Use the Central Limit Theorem to *find* $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$ and *sketch* the sampling distribution of the sample means.
b. *Find the z-scores* that correspond to $\bar{x} = 24.7$ minutes and $\bar{x} = 25.5$ minutes.
c. *Find the cumulative area* that corresponds to each z-score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

Answer: Page A39

EXAMPLE 5

▸ Finding Probabilities for Sampling Distributions

The mean room and board expense per year at four-year colleges is \$7540. You randomly select 9 four-year colleges. What is the probability that the mean room and board is less than \$7800? Assume that the room and board expenses are normally distributed with a standard deviation of \$1245. *(Adapted from 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 \$7540 and a standard deviation of \$415.

$$\mu_{\bar{x}} = \mu = 7540 \quad \text{and} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1245}{\sqrt{9}} = 415$$

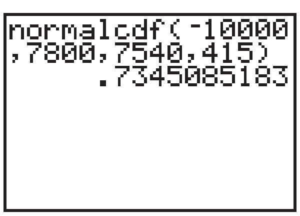
The graph of this distribution is shown at the left. The area to the left of \$7800 is shaded. The z-score that corresponds to \$7800 is

$$z = \frac{7800 - 7540}{1245/\sqrt{9}} = \frac{260}{415} \approx 0.63.$$

So, the probability that the mean room and board expense is less than \$7800 is

$$P(\bar{x} < 7800) = P(z < 0.63)$$
$$= 0.7357.$$

Interpretation So, 73.57% of such samples with $n = 9$ will have a mean less than \$7800 and 26.43% of these sample means will lie outside this interval.



In Example 5, you can use a TI-83/84 Plus to find the probability automatically.

▸ Try It Yourself 5

The average sales price of a single-family house in the United States is \$290,600. You randomly select 12 single-family houses. What is the probability that the mean sales price is more than \$265,000? Assume that the sales prices are normally distributed with a standard deviation of \$36,000. *(Adapted from The U.S. Commerce Department)*

a. Use the Central Limit Theorem to *find* $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$ and *sketch* the sampling distribution of the sample means.
b. *Find the z-score* that corresponds to $\bar{x} = \$265{,}000$.
c. *Find the cumulative area* that corresponds to the z-score and calculate the probability.
d. *Interpret* the results.

Answer: Page A39

The Central Limit Theorem can also be used to investigate unusual events. An unusual event is one that occurs with a probability of less than 5%.

EXAMPLE 6

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

An education finance corporation claims that the average credit card debts carried by undergraduates are normally distributed, with a mean of \$3173 and a standard deviation of \$1120. *(Adapted from Sallie Mae)*

1. What is the probability that a randomly selected undergraduate, who is a credit card holder, has a credit card balance less than \$2700?
2. You randomly select 25 undergraduates who are credit card holders. What is the probability that their mean credit card balance is less than \$2700?
3. Compare the probabilities from (1) and (2) and interpret your answer in terms of the corporation'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 $x = \$2700$ is

$$z = \frac{x - \mu}{\sigma} = \frac{2700 - 3173}{1120} \approx -0.42.$$

So, the probability that the card holder has a balance less than \$2700 is

$$P(x < 2700) = P(z < -0.42) = 0.3372.$$

2. Here, you are asked to find the probability associated with a sample mean $\bar{x}$. The z-score that corresponds to $\bar{x} = \$2700$ is

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{2700 - 3173}{1120/\sqrt{25}} = \frac{-473}{224} \approx -2.11.$$

So, the probability that the mean credit card balance of the 25 card holders is less than \$2700 is

$$P(\bar{x} < 2700) = P(z < -2.11) = 0.0174.$$

3. ***Interpretation*** Although there is about a 34% chance that an undergraduate will have a balance less than \$2700, there is only about a 2% chance that the mean of a sample of 25 will have a balance less than \$2700. Because there is only a 2% chance that the mean of a sample of 25 will have a balance less than \$2700, this is an unusual event. So, it is possible that the corporation's claim that the mean is \$3173 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 liquid crystal display (LCD) computer monitors are normally distributed, with a mean of \$190 and a standard deviation of \$48. (1) What is the probability that a randomly selected LCD computer monitor costs less than \$200? (2) You randomly select 10 LCD computer monitors. What is the probability that their mean cost is less than \$200? (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 A39

5.4 EXERCISES

BUILDING BASIC SKILLS AND VOCABULARY

In Exercises 1–4, a population has a mean $\mu = 150$ and a standard deviation $\sigma = 25$. 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 as a true statement.*

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

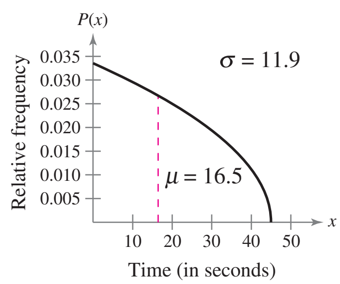
6. As the size of a sample increases, the standard deviation of the distribution of sample means increases.

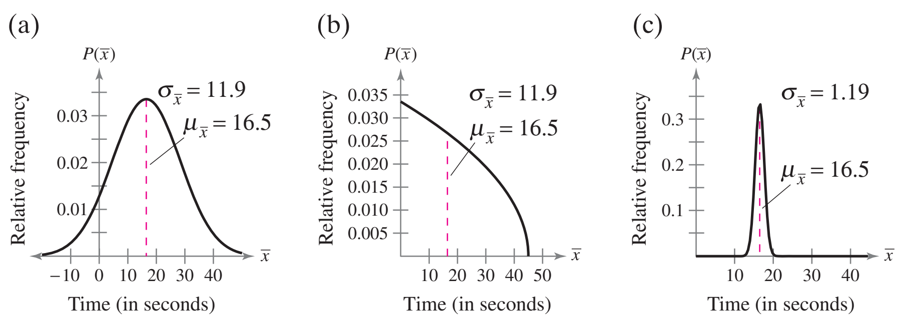
7. A sampling distribution is normal only if the population is normal.

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.

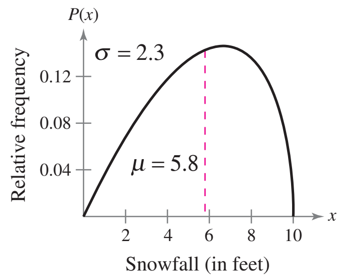
Graphical Analysis *In Exercises 9 and 10, the graph of a population distribution is shown with its mean and standard deviation. Assume that a sample size of 100 is drawn from each population. Decide which of the graphs labeled (a)–(c) would most closely resemble the sampling distribution of the sample means for each graph. Explain your reasoning.*

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

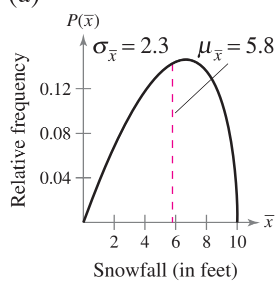




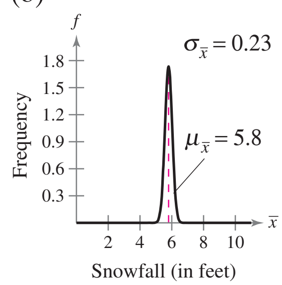
10. The annual snowfall (in feet) for a central New York state county



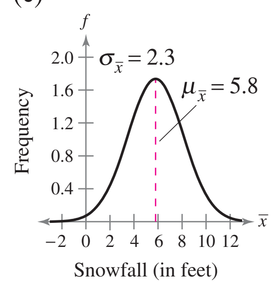
(a)



(b)



(c)



Verifying Properties of Sampling Distributions *In Exercises 11 and 12, 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.*

11. The number of DVDs rented by each of four families in the past month is 8, 4, 16, and 2. Use a sample size of 3.

12. Four friends paid the following amounts for their MP3 players: \$200, \$130, \$270, and \$230. Use a sample size of 2.

Finding Probabilities *In Exercises 13–16, the population mean and standard deviation are given. Find the required probability and determine whether the given sample mean would be considered unusual. If convenient, use technology to find the probability.*

13. For a sample of $n = 64$, find the probability of a sample mean being less than 24.3 if $\mu = 24$ and $\sigma = 1.25$.

14. For a sample of $n = 100$, find the probability of a sample mean being greater than 24.3 if $\mu = 24$ and $\sigma = 1.25$.

15. For a sample of $n = 45$, find the probability of a sample mean being greater than 551 if $\mu = 550$ and $\sigma = 3.7$.

16. For a sample of $n = 36$, find the probability of a sample mean being less than 12,750 or greater than 12,753 if $\mu = 12{,}750$ and $\sigma = 1.7$.

USING AND INTERPRETING CONCEPTS

Using the Central Limit Theorem *In Exercises 17–22, 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.*

17. **Employed Persons** The amounts of time employees at a large corporation work each day are normally distributed, with a mean of 7.6 hours and a standard deviation of 0.35 hour. Random samples of size 12 are drawn from the population and the mean of each sample is determined.

18. **Fly Eggs** The numbers of eggs female house flies lay during their lifetimes are 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.

19. **Photo Printers** The mean price of photo printers on a website is \$235 with a standard deviation of \$62. Random samples of size 20 are drawn from this population and the mean of each sample is determined.

20. **Employees' Ages** The mean age of employees at a large corporation is 47.2 years with 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.

21. **Fresh Vegetables** The per capita consumption of fresh vegetables by people in the United States in a recent year was normally distributed, with a mean of 188.4 pounds and a standard deviation of 54.5 pounds. Random samples of 25 are drawn from this population and the mean of each sample is determined. *(Adapted from U.S. Department of Agriculture)*

22. **Coffee** The per capita consumption of coffee by people in the United States in a recent year was normally distributed, with a mean of 24.2 gallons and a standard deviation of 8.1 gallons. Random samples of 30 are drawn from this population and the mean of each sample is determined. *(Adapted from U.S. Department of Agriculture)*

23. Repeat Exercise 17 for samples of size 24 and 36. What happens to the mean and the standard deviation of the distribution of sample means as the size of the sample increases?

24. Repeat Exercise 18 for samples of size 30 and 45. What happens to the mean and the standard deviation of the distribution of sample means as the size of the sample increases?

Finding Probabilities *In Exercises 25–30, find the probabilities and interpret the results. If convenient, use technology to find the probabilities.*

25. **Salaries** The population mean annual salary for environmental compliance specialists is about \$63,500. A random sample of 35 specialists is drawn from this population. What is the probability that the mean salary of the sample is less than \$60,000? Assume $\sigma = \$6100$. *(Adapted from Salary.com)*

26. **Salaries** The population mean annual salary for flight attendants is \$56,275. A random sample of 48 flight attendants is selected from this population. What is the probability that the mean annual salary of the sample is less than \$56,100? Assume $\sigma = \$1800$. *(Adapted from Salary.com)*

27. **Gas Prices: New England** During a certain week the mean price of gasoline in the New England region was \$2.714 per gallon. A random sample of 32 gas stations is drawn from this population. What is the probability that the mean price for the sample was between \$2.695 and \$2.725 that week? Assume $\sigma = \$0.045$. *(Adapted from U.S. Energy Information Administration)*

28. **Gas Prices: California** During a certain week the mean price of gasoline in California was \$2.999 per gallon. A random sample of 38 gas stations is drawn from this population. What is the probability that the mean price for the sample was between \$3.010 and \$3.025 that week? Assume $\sigma = \$0.049$. *(Adapted from U.S. Energy Information Administration)*

29. Heights of Women The mean height of women in the United States (ages 20–29) is 64.3 inches. A random sample of 60 women in this age group is selected. What is the probability that the mean height for the sample is greater than 66 inches? Assume $\sigma = 2.6$ inches. *(Source: National Center for Health Statistics)*

30. Heights of Men The mean height of men in the United States (ages 20–29) is 69.9 inches. A random sample of 60 men in this age group is selected. What is the probability that the mean height for the sample is greater than 70 inches? Assume $\sigma = 3.0$ inches. *(Source: National Center for Health Statistics)*

31. Which Is More Likely? Assume that the heights given in Exercise 29 are normally distributed. Are you more likely to randomly select 1 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.

32. Which Is More Likely? Assume that the heights given in Exercise 30 are normally distributed. Are you more likely to randomly select 1 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.

33. 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.

34. Make a Decision A machine used to fill half-gallon-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.

35. 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 that 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?

36. 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 that 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?

37. 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 that 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?

38. 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 that 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?

EXTENDING CONCEPTS

39. SAT Scores The mean critical reading SAT score is 501, with a standard deviation of 112. A particular high school claims that its students have unusually high critical reading SAT scores. A random sample of 50 students from this school was selected, and the mean critical reading SAT score was 515. Is the high school justified in its claim? Explain. *(Source: The College Board)*

40. 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 of 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 with 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 $(n/N > 0.05)$, 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 41 and 42, determine if the finite correction factor should be used. If so, use it in your calculations when you find the probability.

41. Gas Prices In a sample of 900 gas stations, the mean price of regular gasoline at the pump was \$2.702 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 \$2.698? *(Adapted from U.S. Department of Energy)*

42. 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)*

Sampling Distribution of Sample Proportions *The sample mean is not the only statistic with a sampling distribution. Every sample statistic, such as the sample median, the sample standard deviation, and the sample proportion, has a sampling distribution. For a random sample of size n, the* ***sample proportion*** *is the number of individuals in the sample with a specified characteristic divided by the sample size. The* ***sampling distribution of sample proportions*** *is the distribution formed when sample proportions of size n are repeatedly taken from a population where the probability of an individual with a specified characteristic is p.*

In Exercises 43–46, suppose three births are randomly selected. There are two equally possible outcomes for each birth, a boy (b) or a girl (g). The number of boys can equal 0, 1, 2, or 3. These correspond to sample proportions of 0, 1/3, 2/3, and 1.

43. List the eight possible samples that can result from randomly selecting three births. For instance, let bbb represent a sample of three boys. Make a table that shows each sample, the number of boys in each sample, and the proportion of boys in each sample.

44. Use the table from Exercise 43 to construct the sampling distribution of the sample proportion of boys from three births. Graph the sampling distribution using a probability histogram. What do you notice about the spread of the histogram as compared to the binomial probability distribution for the number of boys in each sample?

45. Let $x = 1$ represent a boy and $x = 0$ represent a girl. Using these values, find the sample mean for each sample. What do you notice?

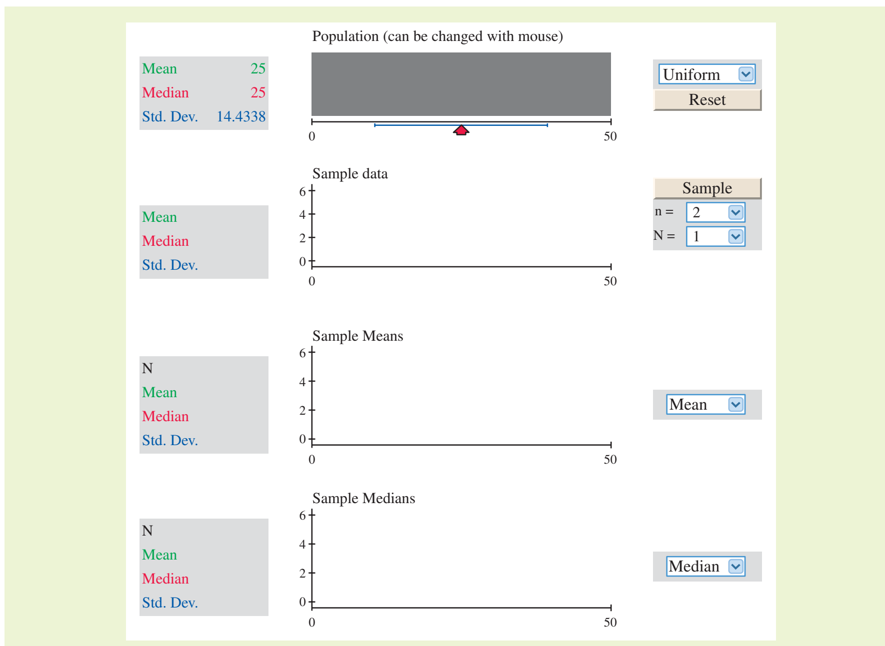
46. Construct a sampling distribution of the sample proportion of boys from four births.

47. Heart Transplants About 77% of all female heart transplant patients will survive for at least 3 years. One hundred five female heart transplant patients are randomly selected. What is the probability that the sample proportion surviving for at least 3 years will be less than 70%? Interpret your results. Assume the sampling distribution of sample proportions is a normal distribution. The mean of the sample proportion is equal to the population proportion p, and the standard deviation is equal to $\sqrt{\frac{pq}{n}}$. *(Source: American Heart Association)*

ACTIVITY 5.4 Sampling Distributions

APPLET

The *sampling distributions* applet allows you to investigate sampling distributions by repeatedly taking samples from a population. The top plot displays the distribution of a population. Several options are available for the population distribution (Uniform, Bell-shaped, Skewed, Binary, and Custom). When SAMPLE is clicked, N random samples of size n will be repeatedly selected from the population. The sample statistics specified in the bottom two plots will be updated for each sample. If N is set to 1 and n is less than or equal to 50, the display will show, in an animated fashion, the points selected from the population dropping into the second plot and the corresponding summary statistic values dropping into the third and fourth plots. Click RESET to stop an animation and clear existing results. Summary statistics for each plot are shown in the panel at the left of the plot.



Explore

Step 1 Specify a distribution.
Step 2 Specify values of n and N.
Step 3 Specify what to display in the bottom two graphs.
Step 4 Click SAMPLE to generate the sampling distributions.

Draw Conclusions

APPLET

1. Run the simulation using $n = 30$ and $N = 10$ for a uniform, a bell-shaped, and a skewed distribution. What is the mean of the sampling distribution of the sample means for each distribution? For each distribution, is this what you would expect?

2. Run the simulation using $n = 50$ and $N = 10$ for a bell-shaped distribution. What is the standard deviation of the sampling distribution of the sample means? According to the formula, what should the standard deviation of the sampling distribution of the sample means be? Is this what you would expect?

5.5 Normal Approximations to Binomial Distributions

WHAT YOU SHOULD LEARN

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

Approximating a Binomial Distribution ▸ Continuity Correction ▸ 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

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

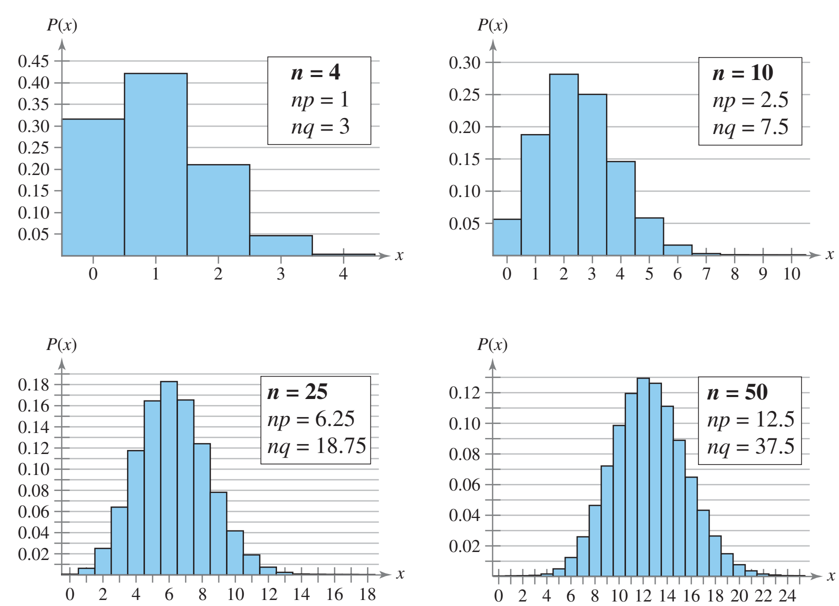
where n is the number of independent trials, p is the probability of success in a single trial, and q is the probability of failure in a single trial.

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 $q = 1 - p$
- p is constant for each trial

To see why this result is valid, look at the following binomial distributions for $p = 0.25$, $q = 1 - 0.25 = 0.75$, and $n = 4$, $n = 10$, $n = 25$, and $n = 50$. Notice that as n increases, the histogram approaches a normal curve.



EXAMPLE 1

Approximating a 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: Opinion Research Corporation)*

1. Sixty-two percent of adults in the United States have an HDTV in their home. You randomly select 45 adults in the United States and ask them if they have an HDTV in their home.

2. Twelve percent of adults in the United States who do not have an HDTV in their home are planning to purchase one in the next two years. You randomly select 30 adults in the United States who do not have an HDTV and ask them if they are planning to purchase one in the next two years.

Solution

1. In this binomial experiment, $n = 45$, $p = 0.62$, and $q = 0.38$. So,

$$np = 45(0.62) = 27.9$$

and

$$nq = 45(0.38) = 17.1.$$

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

$$\mu = np = 27.9$$

and

$$\sigma = \sqrt{npq} = \sqrt{45 \cdot 0.62 \cdot 0.38} \approx 3.26$$

to approximate the distribution of x.

2. In this binomial experiment, $n = 30$, $p = 0.12$, and $q = 0.88$. So,

$$np = 30(0.12) = 3.6$$

and

$$nq = 30(0.88) = 26.4.$$

Because $np < 5$, you cannot use a 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: Opinion Research Corporation)*

Five percent of adults in the United States are planning to purchase a 3D TV in the next two years. You randomly select 125 adults in the United States and ask them if they are planning to purchase a 3D TV in the next two years.

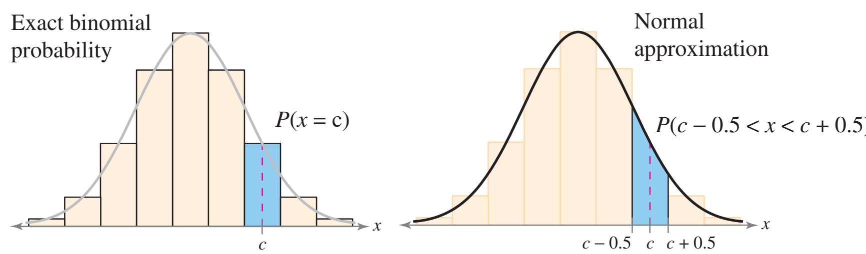
a. *Identify* n, p, and q.
b. *Find* the products np and nq.
c. *Decide* whether you can use a normal distribution to approximate x.
d. *Find* the mean μ and standard deviation σ, if appropriate.

Answer: Page A39

CONTINUITY CORRECTION

A 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 **continuity correction.**



STUDY TIP

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

EXAMPLE 2

Using a Continuity Correction

Use a continuity correction 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 getting at least 158 successes
3. The probability of getting fewer 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 continuity correction to convert each of the following binomial intervals to a normal distribution interval.

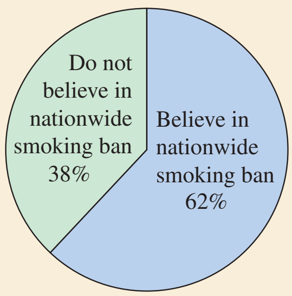
1. The probability of getting between 57 and 83 successes, inclusive
2. The probability of getting at most 54 successes

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

Answer: Page A39

PICTURING THE WORLD

In a survey of U.S. adults, people were asked if there should be a nationwide ban on smoking in all public places. The results of the survey are shown in the following pie chart. (Adapted from Rasmussen Reports)



Assume that this survey is a true indication of the proportion of the population who say there should be a nationwide ban on smoking in all public places. If you sampled 50 adults at random, what is the probability that between 25 and 30, inclusive, would say there should be a nationwide ban on smoking in all public places?

APPROXIMATING BINOMIAL PROBABILITIES

GUIDELINES

Using a Normal Distribution to Approximate Binomial Probabilities

IN WORDS	IN SYMBOLS
1. Verify that a binomial distribution applies.	Specify n, p, and q.
2. Determine if you can use a 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

Sixty-two percent of adults in the United States have an HDTV in their home. You randomly select 45 adults in the United States and ask them if they have an HDTV in their home. What is the probability that fewer than 20 of them respond yes? *(Source: Opinion Research Corporation)*

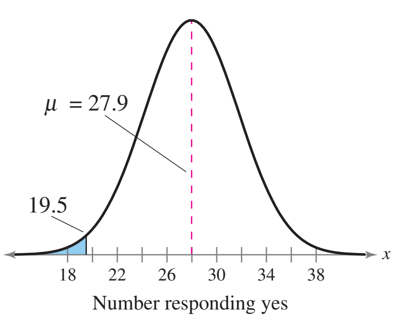
Solution

From Example 1, you know that you can use a normal distribution with $\mu = 27.9$ and $\sigma \approx 3.26$ 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 20" are

$$\ldots 17, 18, 19.$$

To use a normal distribution, add 0.5 to the right-hand boundary 19 to get $x = 19.5$. The graph at the left shows a normal curve with $\mu = 27.9$ and $\sigma \approx 3.26$ and a shaded area to the left of 19.5. The z-score that corresponds to $x = 19.5$ is

$$z = \frac{19.5 - 27.9}{3.26} \approx -2.58.$$



Using the Standard Normal Table,

$$P(z < -2.58) = 0.0049.$$

Interpretation The probability that fewer than 20 people respond yes is approximately 0.0049, or about 0.49%.

Try It Yourself 3

Five percent of adults in the United States are planning to purchase a 3D TV in the next two years. You randomly select 125 adults in the United States and ask them if they are planning to purchase a 3D TV in the next two years. What is the probability that more than 9 respond yes? (See Try It Yourself 1.) *(Source: Opinion Research Corporation)*

a. *Determine* whether you can use a 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 a *continuity correction* to rewrite $P(x > 9)$ 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 A39

EXAMPLE 4

Approximating a Binomial Probability

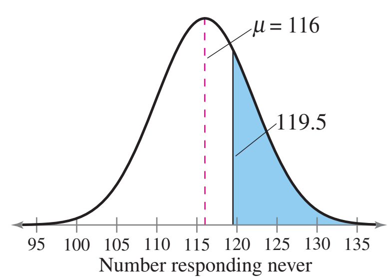
Fifty-eight percent of adults say that they never wear a helmet when riding a bicycle. You randomly select 200 adults in the United States and ask them if they wear a helmet when riding a bicycle. What is the probability that at least 120 adults will say they never wear a helmet when riding a bicycle? *(Source: Consumer Reports National Research Center)*

Solution Because $np = 200 \cdot 0.58 = 116$ and $nq = 200 \cdot 0.42 = 84$, the binomial variable x is approximately normally distributed, with

$$\mu = np = 116 \quad \text{and} \quad \sigma = \sqrt{npq} = \sqrt{200 \cdot 0.58 \cdot 0.42} \approx 6.98.$$

Using the continuity correction, you can rewrite the discrete probability $P(x \geq 120)$ as the continuous probability $P(x \geq 119.5)$. The graph shows a normal curve with $\mu = 116$, $\sigma = 6.98$, and a shaded area to the right of 119.5. The z-score that corresponds to 119.5 is

$$z = \frac{119.5 - 116}{6.98} \approx 0.50.$$



So, the probability that at least 120 will say yes is approximately

$$P(x \geq 119.5) = P(z \geq 0.50)$$
$$= 1 - P(z \leq 0.50) = 1 - 0.6915 = 0.3085.$$

Try It Yourself 4

In Example 4, what is the probability that at most 100 adults will say they never wear a helmet when riding a bicycle?

a. *Determine* whether you can use a normal distribution to approximate the binomial variable (see Example 4).
b. Find the *mean* μ and the *standard deviation* σ for the distribution (see Example 4).
c. Apply a *continuity correction* to rewrite $P(x \leq 100)$ 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 A39

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 0. 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 0.

```
normalcdf(119.5,
10000,116,6.98)
          .3080333206
```

In Example 4, you can use a TI-83/84 Plus to find the probability once the mean, standard deviation, and continuity correction are calculated. Use 10,000 for the upper bound.

EXAMPLE 5

Approximating a Binomial Probability

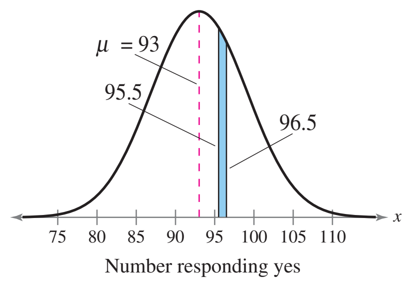
A survey reports that 62% of Internet users use Windows® Internet Explorer® as their browser. You randomly select 150 Internet users and ask them whether they use Internet Explorer® as their browser. What is the probability that exactly 96 will say yes? *(Source: Net Applications)*

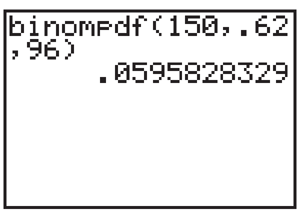
Solution

Because $np = 150 \cdot 0.62 = 93$ and $nq = 150 \cdot 0.38 = 57$, the binomial variable x is approximately normally distributed, with

$$\mu = np = 93 \quad \text{and} \quad \sigma = \sqrt{npq} = \sqrt{150 \cdot 0.62 \cdot 0.38} \approx 5.94.$$

Using the continuity correction, you can rewrite the discrete probability $P(x = 96)$ as the continuous probability $P(95.5 < x < 96.5)$. The graph shows a normal curve with $\mu = 93$, $\sigma = 5.94$, and a shaded area between 95.5 and 96.5.





The approximation in Example 5 is almost exactly equal to the exact probability found using the binompdf(command on a TI-83/84 Plus.

The z-scores that correspond to 95.5 and 96.5 are

$$z_1 = \frac{95.5 - 93}{5.94} \approx 0.42 \quad \text{and} \quad z_2 = \frac{96.5 - 93}{5.94} \approx 0.59.$$

So, the probability that exactly 96 Internet users will say they use Internet Explorer® is

$$\begin{aligned} P(95.5 < x < 96.5) &= P(0.42 < z < 0.59) \\ &= P(z < 0.59) - P(z < 0.42) \\ &= 0.7224 - 0.6628 \\ &= 0.0596. \end{aligned}$$

Interpretation The probability that exactly 96 of the Internet users will say they use Internet Explorer® is approximately 0.0596, or about 6%.

Try It Yourself 5

A survey reports that 24% of Internet users use Mozilla® Firefox® as their browser. You randomly select 150 Internet users and ask them whether they use Firefox® as their browser. What is the probability that exactly 27 will say yes? *(Source: Net Applications)*

a. *Determine* whether you can use a normal distribution to approximate the binomial variable.
b. Find the *mean* μ and the *standard deviation* σ for the distribution.
c. Apply a *continuity correction* to rewrite $P(x = 27)$ 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 A39

5.5 EXERCISES

BUILDING BASIC SKILLS AND VOCABULARY

1. What are the properties of a binomial experiment?

2. What are the conditions for using a normal distribution to approximate a binomial distribution?

In Exercises 3–6, the sample size n, probability of success p, and probability of failure q are given for a binomial experiment. Decide whether you can use a normal distribution to approximate the random variable x.

3. $n = 24,\ p = 0.85,\ q = 0.15$

4. $n = 15,\ p = 0.70,\ q = 0.30$

5. $n = 18,\ p = 0.90,\ q = 0.10$

6. $n = 20,\ p = 0.65,\ q = 0.35$

Approximating a Binomial Distribution *In Exercises 7–12, a binomial experiment is given. Decide whether you can use a normal distribution to approximate the binomial distribution. If you can, find the mean and standard deviation. If you cannot, explain why.*

7. House Contract A survey of U.S. adults found that 85% read every word or at least enough to understand a contract for buying or selling a home before signing. You randomly select 10 adults and ask them if they read every word or at least enough to understand a contract for buying or selling a home before signing. *(Source: FindLaw.com)*

8. 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 them if they would want their organs transplanted into a patient who needs them if they were killed in an accident. *(Source: USA Today)*

9. Multivitamins A survey of U.S. adults found that 55% have used a multivitamin in the past 12 months. You randomly select 50 adults and ask them if they have used a multivitamin in the past 12 months. *(Source: Harris Interactive)*

10. Happiness at Work A survey of U.S. adults found that 19% are happy with their current employer. You randomly select 30 adults and ask them if they are happy with their current employer. *(Source: Opinion Research Corporation)*

11. Going Green A survey of U.S. adults found that 76% would pay more for an environmentally friendly product. You randomly select 20 adults and ask them if they would pay more for an environmentally friendly product. *(Source: Opinion Research Corporation)*

12. Online Habits A survey of U.S. adults found that 61% look online for health information. You randomly select 15 adults and ask them if they look online for health information. *(Source: Pew Research Center)*

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

Binomial Probability	Normal Probability
13. $P(x > 109)$	(a) $P(x > 109.5)$
14. $P(x \geq 109)$	(b) $P(x < 108.5)$
15. $P(x \leq 109)$	(c) $P(x \leq 109.5)$
16. $P(x < 109)$	(d) $P(x \geq 108.5)$

In Exercises 17–22, a binomial probability is given. Write the probability in words. Then, use a continuity correction to convert the binomial probability to a normal distribution probability.

17. $P(x < 25)$ **18.** $P(x \geq 110)$ **19.** $P(x = 33)$

20. $P(x > 65)$ **21.** $P(x \leq 150)$ **22.** $P(55 < x < 60)$

USING AND INTERPRETING CONCEPTS

Approximating Binomial Probabilities *In Exercises 23–30, decide whether you can use a 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 a binomial distribution to find the indicated probabilities.*

23. Internet Use A survey of U.S. adults ages 18–29 found that 93% use the Internet. You randomly select 100 adults ages 18–29 and ask them if they use the Internet. *(Source: Pew Research Center)*

(a) Find the probability that exactly 90 people say they use the Internet.
(b) Find the probability that at least 90 people say they use the Internet.
(c) Find the probability that fewer than 90 people say they use the Internet.
(d) Are any of the probabilities in parts (a)–(c) unusual? Explain.

24. Internet Use A survey of U.S. adults ages 50–64 found that 70% use the Internet. You randomly select 80 adults ages 50–64 and ask them if they use the Internet. *(Source: Pew Research Center)*

(a) Find the probability that at least 70 people say they use the Internet.
(b) Find the probability that exactly 50 people say they use the Internet.
(c) Find the probability that more than 60 people say they use the Internet.
(d) Are any of the probabilities in parts (a)–(c) unusual? Explain.

25. Favorite Sport A survey of U.S. adults found that 35% say their favorite sport is professional football. You randomly select 150 adults and ask them if their favorite sport is professional football. *(Source: Harris Interactive)*

(a) Find the probability that at most 75 people say their favorite sport is professional football.
(b) Find the probability that more than 40 people say their favorite sport is professional football.
(c) Find the probability that between 50 and 60 people, inclusive, say their favorite sport is professional football.
(d) Are any of the probabilities in parts (a)–(c) unusual? Explain.

26. College Graduates About 34% of workers in the United States are college graduates. You randomly select 50 workers and ask them if they are a college graduate. *(Source: U.S. Bureau of Labor Statistics)*

(a) Find the probability that exactly 12 workers are college graduates.
(b) Find the probability that more than 23 workers are college graduates.
(c) Find the probability that at most 18 workers are college graduates.
(d) A committee is looking for 30 working college graduates to volunteer at a career fair. The committee randomly selects 125 workers. What is the probability that there will not be enough college graduates?

27. Public Transportation Five percent of workers in the United States use public transportation to get to work. You randomly select 250 workers and ask them if they use public transportation to get to work. *(Source: U.S. Census Bureau)*

(a) Find the probability that exactly 16 workers will say yes.

(b) Find the probability that at least 9 workers will say yes.

(c) Find the probability that fewer than 16 workers will say yes.

(d) A transit authority offers discount rates to companies that have at least 30 employees who use public transportation to get to work. There are 500 employees in a company. What is the probability that the company will not get the discount?

28. Concert Tickets A survey of U.S. adults who attend at least one music concert a year found that 67% say concert tickets are too expensive. You randomly select 12 adults who attend at least one music concert a year and ask them if concert tickets are too expensive. *(Source: Rasmussen Reports)*

(a) Find the probability that fewer than 4 people say that concert tickets are too expensive.

(b) Find the probability that between 7 and 9 people, inclusive, say that concert tickets are too expensive.

(c) Find the probability that at most 10 people say that concert tickets are too expensive.

(d) Are any of the probabilities in parts (a)–(c) unusual? Explain.

29. News A survey of U.S. adults ages 18–24 found that 34% get no news on an average day. You randomly select 200 adults ages 18–24 and ask them if they get news on an average day. *(Source: Pew Research Center)*

(a) Find the probability that at least 85 people say they get no news on an average day.

(b) Find the probability that fewer than 66 people say they get no news on an average day.

(c) Find the probability that exactly 68 people say they get no news on an average day.

(d) A college English teacher wants students to discuss current events. The teacher randomly selects six students from the class. What is the probability that none of the students can talk about current events because they get no news on an average day.

30. Long Work Weeks A survey of U.S. workers found that 2.9% work more than 70 hours per week. You randomly select 10 workers in the United States and ask them if they work more than 70 hours per week.

(a) Find the probability that at most 3 people say they work more than 70 hours per week.

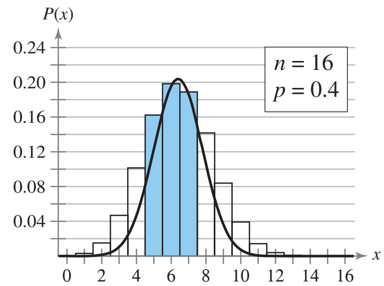
(b) Find the probability that at least 1 person says he or she works more than 70 hours per week.

(c) Find the probability that more than 2 people say they work more than 70 hours per week.

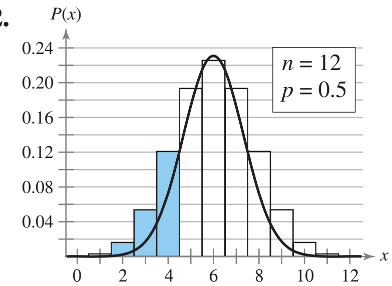
(d) A large company is concerned about overworked employees who work more than 70 hours per week. The company randomly selects 50 employees. What is the probability there will be no employee working more than 70 hours?

Graphical Analysis *In Exercises 31 and 32, 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.*

31.

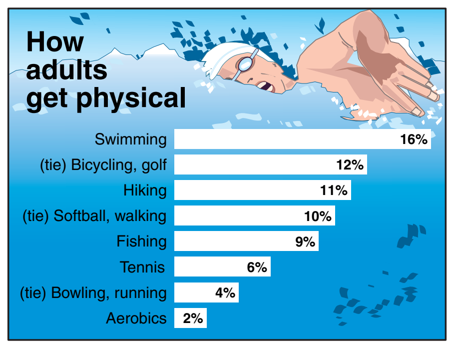


32.



EXTENDING CONCEPTS

Getting Physical *In Exercises 33 and 34, 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 participated in at least one sport, and they gave their favorite sport.*



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

34. You randomly select 300 people in the United States ages 33 to 51 and ask them if they regularly participate in at least one sport. Of the 200 who say yes, 9% say they participate in hiking. How likely is this result? Do you think this sample is a good one? Explain your reasoning.

Testing a Drug *In Exercises 35 and 36, 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, this claim will be accepted.*

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

36. 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

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 approximately 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

Unusual Events 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 115 or more. In fact, this will happen almost 16% of the time. It *would* be, however, highly unusual to take random samples of 100 values from that population and obtain a sample with a mean of 115 or more. Because the population is normally distributed, the mean of the sample distribution will be 100, and the standard deviation will be 1.5. A sample mean of 115 lies 10 standard deviations above the mean. This would be an extremely unusual event. When an event this unusual occurs, it is a good idea to question the original claimed value of the mean.

Although normal distributions are common in many populations, people try to make *non-normal* statistics fit a normal distribution. The statistics used for normal distributions are often inappropriate when the distribution is obviously non-normal.

EXERCISES

1. ***Is It Unusual?*** A population is normally distributed, with a mean of 100 and a standard deviation of 15. Determine if either of the following events is unusual. Explain your reasoning.
   a. The mean of a sample of 3 is 115 or more.
   b. The mean of a sample of 20 is 105 or more.
2. ***Find the Error*** The mean age of students at a high school is 16.5, with a standard deviation of 0.7. You use the Standard Normal Table to help you determine that the probability of selecting one student at random and finding his or her age to be more than 17.5 years is about 8%. What is the error in this problem?
3. Give an example of a distribution that might be non-normal.

5 CHAPTER SUMMARY

What did you learn?	EXAMPLE(S)	REVIEW EXERCISES
Section 5.1		
■ How to interpret graphs of normal probability distributions	*1, 2*	*1–6*
■ How to find areas under the standard normal curve	*3–6*	*7–28*
Section 5.2		
■ How to find probabilities for normally distributed variables	*1–3*	*29–38*
Section 5.3		
■ How to find a z-score given the area under the normal curve	*1, 2*	*39–46*
■ How to transform a z-score to an x-value $x = \mu + z\sigma$	*3*	*47, 48*
■ How to find a specific data value of a normal distribution given the probability	*4, 5*	*49–52*
Section 5.4		
■ How to find sampling distributions and verify their properties	*1*	*53, 54*
■ How to interpret the Central Limit Theorem $\mu_{\bar{x}} = \mu$ Mean $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ Standard deviation	*2, 3*	*55, 56*
■ How to apply the Central Limit Theorem to find the probability of a sample mean	*4–6*	*57–62*
Section 5.5		
■ How to decide when a normal distribution can approximate a binomial distribution $\mu = np$ Mean $\sigma = \sqrt{npq}$ Standard deviation	*1*	*63, 64*
■ How to find the continuity correction	*2*	*65–68*
■ How to use a normal distribution to approximate binomial probabilities	*3–5*	*69, 70*

5 REVIEW EXERCISES

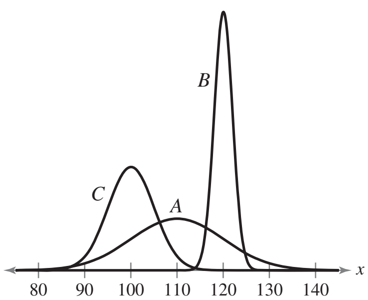
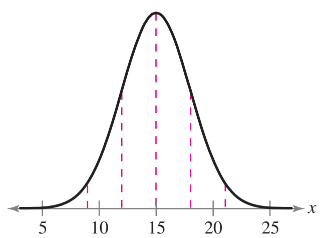


FIGURE FOR EXERCISES 3 AND 4

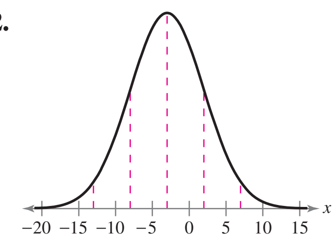
SECTION 5.1

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

1.



2.



In Exercises 3 and 4, use the normal curves shown.

3. Which normal curve has the greatest mean? Explain your reasoning.

4. Which normal curve has the greatest standard deviation? Explain your reasoning.

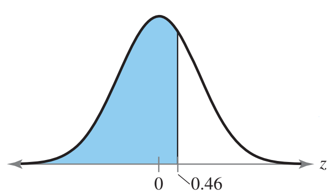
In Exercises 5 and 6, use the following information and standard scores to investigate observations about a normal population. A batch of 2500 resistors is normally distributed, with a mean resistance of 1.5 ohms and a standard deviation of 0.08 ohm. Four resistors are randomly selected and tested. Their resistances are measured at 1.32, 1.54, 1.66, and 1.78 ohms.

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

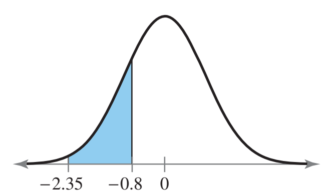
6. Are there any unusual observations?

In Exercises 7 and 8, find the area of the indicated region under the standard normal curve. If convenient, use technology to find the area.

7.



8.



In Exercises 9–20, find the indicated area under the standard normal curve. If convenient, use technology to find the area.

9. To the left of $z = 0.33$

10. To the left of $z = -1.95$

11. To the right of $z = -0.57$

12. To the right of $z = 3.22$

13. To the left of $z = -2.825$

14. To the right of $z = 0.015$

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

16. Between $z = -1.55$ and $z = 1.04$

17. Between $z = 0.05$ and $z = 1.71$

18. Between $z = -2.68$ and $z = 2.68$

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

20. To the left of $z = 0.64$ and to the right of $z = 3.415$

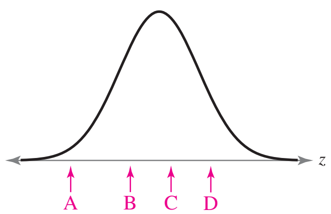


FIGURE FOR EXERCISES 21 AND 22

In Exercises 21 and 22, use the following information. In a recent year, the ACT scores for the reading portion of the test were normally distributed, with a mean of 21.4 and a standard deviation of 6.2. The test scores of four students selected at random are 17, 29, 8, and 23. (Source: ACT, Inc.)

21. Without converting to z-scores, match the values with the letters A, B, C, and D on the given graph.

22. Find the z-score that corresponds to each value and check your answers in Exercise 21. Are any of the values unusual? Explain.

In Exercises 23–28, find the indicated probabilities. If convenient, use technology to find the probability.

23. $P(z < 1.28)$

24. $P(z > -0.74)$

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

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

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

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

SECTION 5.2

In Exercises 29–34, assume the random variable x is normally distributed, with mean $\mu = 74$ and standard deviation $\sigma = 8$. Find the indicated probability.

29. $P(x < 84)$

30. $P(x < 55)$

31. $P(x > 80)$

32. $P(x > 71.6)$

33. $P(60 < x < 70)$

34. $P(72 < x < 82)$

In Exercises 35 and 36, find the indicated probabilities.

35. 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)

36. 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)

37. Can any of the events in Exercise 35 be considered unusual? Explain your reasoning.

38. Can any of the events in Exercise 36 be considered unusual? Explain your reasoning.

SECTION 5.3

In Exercises 39–44, 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 convenient, use technology to find the z-score.

39. 0.4721	**40.** 0.1	**41.** 0.8708
42. P_2	**43.** P_{85}	**44.** P_{46}

45. Find the z-score that has 30.5% of the distribution's area to its right.

46. Find the z-score for which 94% of the distribution's area lies between $-z$ and z.

Braking Distance of a Cadillac Escalade

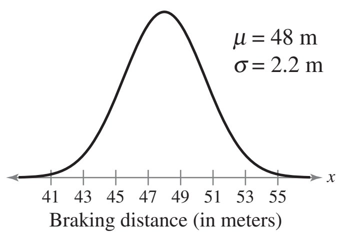


FIGURE FOR EXERCISES 47–52

In Exercises 47–52, use the following information. On a dry surface, the braking distance (in meters) of a Cadillac Escalade can be approximated by a normal distribution, as shown in the graph at the left. (Adapted from Consumer Reports)

47. Find the braking distance of a Cadillac Escalade that corresponds to $z = -2.5$.

48. Find the braking distance of a Cadillac Escalade that corresponds to $z = 1.2$.

49. What braking distance of a Cadillac Escalade represents the 95th percentile?

50. What braking distance of a Cadillac Escalade represents the third quartile?

51. What is the shortest braking distance of a Cadillac Escalade that can be in the top 10% of braking distances?

52. What is the longest braking distance of a Cadillac Escalade that can be in the bottom 5% of braking distances?

SECTION 5.4

In Exercises 53 and 54, use the given population to find the mean and standard deviation of the population and the mean and standard deviation of the sampling distribution. Compare the values.

53. A corporation has four executives. The number of minutes of overtime per week reported by each is 90, 120, 160, and 210. Draw three executives' names from this population, with replacement.

54. There are four residents sharing a house. The number of times each washes their car each month is 1, 2, 0, and 3. Draw two names from this population, with replacement.

In Exercises 55 and 56, 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.

55. The per capita consumption of citrus fruits by people in the United States in a recent year was normally distributed, with a mean of 76.0 pounds and a standard deviation of 20.5 pounds. Random samples of 35 people are drawn from this population and the mean of each sample is determined. *(Adapted from U.S. Department of Agriculture)*

56. The per capita consumption of red meat by people in the United States in a recent year was normally distributed, with a mean of 108.3 pounds and a standard deviation of 35.1 pounds. Random samples of 40 people are drawn from this population and the mean of each sample is determined. *(Adapted from U.S. Department of Agriculture)*

In Exercises 57–62, find the probabilities for the sampling distributions. Interpret the results.

57. Refer to Exercise 35. 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 35.

58. Refer to Exercise 36. 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 36.

59. The mean annual salary for chauffeurs is \$29,200. A sample of 45 chauffeurs is randomly selected. What is the probability that the mean annual salary is (a) less than \$29,000 and (b) more than \$31,000? Assume $\sigma = \$1500$. *(Source: Salary.com)*

60. The mean value of land and buildings per acre for farms is \$1300. A sample of 36 farms is randomly selected. 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 $\sigma = \$250$.

61. 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?

62. Mean rent in a city is \$500 per month with a standard deviation of \$30. The rents are normally distributed. You randomly select 15 apartments in this city. What is the probability that the mean rent will be more than \$525?

SECTION 5.5

In Exercises 63 and 64, a binomial experiment is given. Decide whether you can use a normal distribution to approximate the binomial distribution. If you can, find the mean and standard deviation. If you cannot, explain why.

63. In a recent year, the American Cancer Society said that the five-year survival rate for new cases of stage 1 kidney cancer is 96%. You randomly select 12 men who were new stage 1 kidney cancer cases this year and calculate the five-year survival rate of each. *(Source: American Cancer Society, Inc.)*

64. A survey indicates that 75% of U.S. adults who go to the theater at least once a month think movie tickets are too expensive. You randomly select 30 adults and ask them if they think movie tickets are too expensive. *(Source: Rasmussen Reports)*

In Exercises 65–68, write the binomial probability as a normal probability using the continuity correction.

65. $P(x \geq 25)$

66. $P(x \leq 36)$

67. $P(x = 45)$

68. $P(x = 50)$

In Exercises 69 and 70, decide whether you can use a 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 a binomial distribution to find the indicated probabilities.

69. Seventy 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 them if they keep at least part of their savings in a savings account. Find the probability that at most 20 children will say yes. *(Source: International Communications Research for Merrill Lynch)*

70. Thirty-one percent of people in the United States have type A^+ blood. You randomly select 15 people in the United States and ask them if their blood type is A^+. Find the probability that more than 8 adults say they have A^+ blood. *(Source: American Association of Blood Banks)*

5 CHAPTER QUIZ

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) $P(z > -2.54)$
   (b) $P(z < 3.09)$
   (c) $P(-0.88 < z < 0.88)$
   (d) $P(z < -1.445 \text{ or } z > -0.715)$
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. Students taking a standardized IQ test had a mean score of 100 with a standard deviation of 15. Assume that the scores are normally distributed. *(Adapted from Audiblox)*

3. Find the probability that a student had a score higher than 125. Is this an unusual event? Explain.
4. Find the probability that a student had a score between 95 and 105. Is this an unusual event? Explain.
5. What percent of the students had an IQ score that is greater than 112?
6. If 2000 students are randomly selected, how many would be expected to have an IQ score that is less than 90?
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 10% of the scores?
9. A random sample of 60 students is drawn from this population. What is the probability that the mean IQ score is greater than 105? Interpret your result.
10. Are you more likely to randomly select one student with an IQ score greater than 105 or are you more likely to randomly select a sample of 15 students with a mean IQ score greater than 105? Explain.

In Exercises 11 and 12, use the following information. In a survey of adults under age 65, 81% say they are concerned about the amount and security of personal online data that can be accessed by cybercriminals and hackers. You randomly select 35 adults and ask them if they are concerned about the amount and security of personal online data that can be accessed by cybercriminals and hackers. *(Source: Financial Times/Harris Poll)*

11. Decide whether you can use a 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 20 adults say they are concerned about the amount and security of personal online data that can be accessed by cybercriminals and hackers. Interpret the result.

PUTTING IT ALL TOGETHER

Real Statistics — Real Decisions

You are the human resources director for a corporation and want to implement a health improvement program for employees to decrease employee medical absences. You perform a six-month study with a random sample of employees. Your goal is to decrease absences by 50%. (Assume all data are normally distributed.)

EXERCISES

1. ***Preliminary Thoughts***

You got the idea for this health improvement program from a national survey in which 75% of people who responded said they would participate in such a program if offered by their employer. You randomly select 60 employees and ask them whether they would participate in such a program.

(a) Find the probability that exactly 35 will say yes.

(b) Find the probability that at least 40 will say yes.

(c) Find the probability that fewer than 20 will say yes.

(d) Based on the results in parts (a)–(c), explain why you chose to perform the study.

2. ***Before the Program***

Before the study, the mean number of absences during a six-month period of the participants was 6, with a standard deviation of 1.5. An employee is randomly selected.

(a) Find the probability that the employee's number of absences is less than 5.

(b) Find the probability that the employee's number of absences is between 5 and 7.

(c) Find the probability that the employee's number of absences is more than 7.

3. ***After the Program***

The graph at the right represents the results of the study.

(a) What is the mean number of absences for employees? Explain how you know.

(b) Based on the results, was the goal of decreasing absences by 50% reached?

(c) Describe how you would present your results to the board of directors of the corporation.

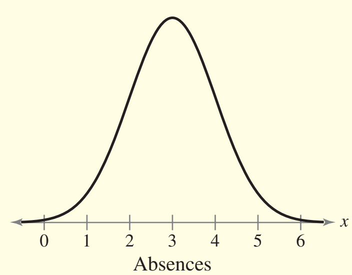


FIGURE FOR EXERCISE 3

TECHNOLOGY

MINITAB | EXCEL | TI-83/84 PLUS

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 2009 is shown below.

Age Distribution in the U.S.

Class	Class midpoint	Relative frequency
0–4	2	6.9%
5–9	7	6.6%
10–14	12	6.6%
15–19	17	7.1%
20–24	22	6.9%
25–29	27	7.0%
30–34	32	6.4%
35–39	37	6.9%
40–44	42	7.1%
45–49	47	7.5%
50–54	52	7.1%
55–59	57	6.1%
60–64	62	5.0%
65–69	67	3.7%
70–74	72	2.9%
75–79	77	2.4%
80–84	82	1.9%
85–89	87	1.2%
90–94	92	0.5%
95–99	97	0.2%

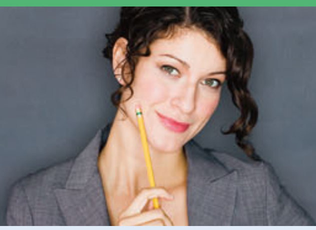
EXERCISES

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.

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-83/84 Plus.



CUMULATIVE REVIEW

Chapters 3 – 5

1. A survey of voters in the United States found that 15% rate the U.S. health care system as excellent. You randomly select 50 voters and ask them how they rate the U.S. health care system. *(Source: Rasmussen Reports)*

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

(b) Find the probability that at most 14 voters rate the U.S. health care system as excellent.

(c) Is it unusual for 14 out of 50 voters to rate the U.S. health care system as excellent? Explain your reasoning.

In Exercises 2 and 3, use the probability distribution to find the (a) mean, (b) variance, (c) standard deviation, and (d) expected value of the probability distribution, and (e) interpret the results.

2. The table shows the distribution of family household sizes in the United States for a recent year. *(Source: U.S. Census Bureau)*

x	2	3	4	5	6	7
$P(x)$	0.427	0.227	0.200	0.093	0.034	0.018

3. The table shows the distribution of fouls per game for a player in a recent NBA season. *(Source: NBA.com)*

x	0	1	2	3	4	5	6
$P(x)$	0.012	0.049	0.159	0.256	0.244	0.195	0.085

4. Use the probability distribution in Exercise 3 to find the probability of randomly selecting a game in which the player had (a) fewer than four fouls, (b) at least three fouls, and (c) between two and four fouls, inclusive.

5. From a pool of 16 candidates, 9 men and 7 women, the offices of president, vice president, secretary, and treasurer will be filled. (a) In how many different ways can the offices be filled? (b) What is the probability that all four of the offices are filled by women?

In Exercises 6–11, use the Standard Normal Table to find the indicated area under the standard normal curve.

6. To the left of $z = 0.72$

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

8. To the right of $z = -0.84$

9. Between $z = 0$ and $z = 2.95$

10. Between $z = -1.22$ and $z = -0.26$

11. To the left of $z = 0.12$ or to the right of $z = 1.72$

12. Forty-five percent of adults say they are interested in regularly measuring their carbon footprint. You randomly select 11 adults and ask them if they are interested in regularly measuring their carbon footprint. Find the probability that the number of adults who say they are interested is (a) exactly eight, (b) at least five, and (c) less than two. Are any of these events unusual? Explain your reasoning. *(Source: Sacred Heart University Polling)*

13. An auto parts seller finds that 1 in every 200 parts sold is defective. Use the geometric distribution to find the probability that (a) the first defective part is the tenth part sold, (b) the first defective part is the first, second, or third part sold, and (c) none of the first 10 parts sold are defective.

14. The table shows the results of a survey in which 2,944,100 public and 401,900 private school teachers were asked about their full-time teaching experience. *(Adapted from U.S. National Center for Education Statistics)*

	Public	Private	Total
Less than 3 years	177,300	27,600	204,900
3 to 9 years	995,800	154,500	1,150,300
10 to 20 years	906,300	111,600	1,017,900
20 years or more	864,700	108,200	972,900
Total	2,944,100	401,900	3,346,000

(a) Find the probability that a randomly selected private school teacher has 10 to 20 years of full-time teaching experience.

(b) Given that a randomly selected teacher has 3 to 9 years of full-time experience, find the probability that the teacher is at a public school.

(c) Are the events "being a public school teacher" and "having 20 years or more of full-time teaching experience" independent? Explain.

(d) Find the probability that a randomly selected teacher is either at a public school or has less than 3 years of full-time teaching experience.

(e) Find the probability that a randomly selected teacher has 3 to 9 years of full-time teaching experience or is at a private school.

15. The initial pressures for bicycle tires when first filled are normally distributed, with a mean of 70 pounds per square inch (psi) and a standard deviation of 1.2 psi.

(a) Random samples of size 40 are drawn from this population and the mean of each sample is determined. Use the Central Limit Theorem to find the mean and standard error of the mean of the sampling distribution. Then sketch a graph of the sampling distribution of sample means.

(b) A random sample of 15 tires is drawn from this population. What is the probability that the mean tire pressure of the sample $\overline{x}$ is less than 69 psi?

16. The life spans of car batteries are normally distributed, with a mean of 44 months and a standard deviation of 5 months.

(a) A car battery is selected at random. Find the probability that the life span of the battery is less than 36 months.

(b) A car battery is selected at random. Find the probability that the life span of the battery is between 42 and 60 months.

(c) What is the shortest life expectancy a car battery can have and still be in the top 5% of life expectancies?

17. A florist has 12 different flowers from which floral arrangements can be made. (a) If a centerpiece is to be made using four different flowers, how many different centerpieces can be made? (b) What is the probability that the four flowers in the centerpiece are roses, gerbers, hydrangeas, and callas?

18. About fifty percent of adults say they feel vulnerable to identity theft. You randomly select 16 adults and ask them if they feel vulnerable to identity theft. Find the probability that the number who say they feel vulnerable is (a) exactly 12, (b) no more than 6, and (c) more than 7. Are any of these events unusual? Explain your reasoning. *(Adapted from KRC Research for Fellowes)*