In Chapter 6, we introduced discrete probability distributions and, in particular, the binomial probability distribution. We computed probabilities for this discrete distribution using its probability distribution function. However, we could also determine the probability of any discrete random variable from its probability histogram. For example, the figure shows the probability histogram for the binomial random variable $X$ with $n = 5$ and $p = 0.35$. From the probability histogram, we can see $P(1) = 0.31$. Notice that the width of each rectangle in the probability histogram is 1. Since the area of a rectangle equals height times width, we can think of $P(1)$ as the area of the rectangle corresponding to $x = 1$. Thinking of probability in this fashion makes the transition from computing discrete probabilities to continuous probabilities much easier.

In this chapter, we discuss two continuous distributions, the uniform distribution and the normal distribution. The greater part of the discussion will focus on the normal distribution, which has many uses and applications.
7.1 PROPERTIES OF THE NORMAL DISTRIBUTION

Preparing for This Section  Before getting started, review the following:

• Continuous variable (Section 1.1, p. 8)  
• The Empirical Rule (Section 3.2, pp. 150–152)  
• z-score (Section 3.4, pp. 167–168)  
• Rules for a discrete probability distribution (Section 6.1, p. 332)

Objectives

1. Utilize the uniform probability distribution
2. Graph a normal curve
3. State the properties of the normal curve
4. Explain the role of area in the normal density function
5. Describe the relation between a normal random variable and a standard normal random variable

1 Utilize the Uniform Probability Distribution

We illustrate a uniform distribution using an example. Using the uniform distribution makes it easy to see the relation between area and probability.

EXAMPLE 1  The Uniform Distribution

Imagine that a friend of yours is always late. Let the random variable \( X \) represent the time from when you are supposed to meet your friend until he shows up. Further suppose that your friend could be on time or up to 30 minutes late with all intervals of equal time between and being equally likely. For example, your friend is just as likely to be from 3 to 4 minutes late as he is to be 25 to 26 minutes late. The random variable \( X \) can be any value in the interval \( 0 \leq X \leq 30 \).

When we compute probabilities for discrete random variables, we usually substitute the value of the random variable into a formula.

Things are not as easy for continuous random variables. Since an infinite number of outcomes are possible for continuous random variables, the probability of observing a particular value of a continuous random variable is zero. For example, the probability that your friend is exactly 12.9438823 minutes late is zero. This result is based on the fact that classical probability is found by dividing the number of ways an event can occur by the total number of possibilities. There is one way to observe 12.9438823, and there are an infinite number of possible values between 0 and 30, so we get a probability that is zero. To resolve this problem, we compute probabilities of continuous random variables over an interval of values. For example, we might compute the probability that your friend is between 10 and 15 minutes late. To find probabilities for continuous random variables, we use probability density functions.

A probability density function (pdf) is an equation used to compute probabilities of continuous random variables. It must satisfy the following two properties:

1. The total area under the graph of the equation over all possible values of the random variable must equal 1.
2. The height of the graph of the equation must be greater than or equal to 0 for all possible values of the random variable. That is, the graph of the equation must lie on or above the horizontal axis for all possible values of the random variable.
The area under the graph of a density function over an interval represents the probability of observing a value of the random variable in that interval.

Property 1 is similar to the rule for discrete probability distributions that stated the sum of the probabilities must add up to 1. Property 2 is similar to the rule that stated all probabilities must be greater than or equal to 0.

Figure 1 illustrates the properties for the example about your friend who is always late. Since all possible values of the random variable between 0 and 30 are equally likely, the graph of the probability density function for uniform random variables is a rectangle. Because the random variable is any number between 0 and 30 inclusive, the width of the rectangle is 30. Since the area under the graph of the probability density function must equal 1, and the area of a rectangle equals height times width, the height of the rectangle must be $\frac{1}{30}$.

In Other Words
To find probabilities for continuous random variables, we do not use probability distribution functions (as we did for discrete random variables). Instead, we use probability density functions. The word density is used because it refers to the number of individuals per unit of area.

A pressing question remains: How do we use density functions to find probabilities of continuous random variables?

The area under the graph of a density function over an interval represents the probability of observing a value of the random variable in that interval.

The following example illustrates this statement.

**EXAMPLE 2**

**Area as a Probability**

**Problem:** Refer to the situation presented in Example 1. What is the probability that your friend will be between 10 and 20 minutes late the next time you meet him?

**Approach:** Figure 1 presented the graph of the density function. We need to find the area under the graph between 10 and 20 minutes.

**Solution:** Figure 2 presents the graph of the density function with the area we wish to find shaded in green.

The width of the shaded rectangle is 10 and its height is $\frac{1}{30}$. Therefore, the area between 10 and 20 is $10 \times \left( \frac{1}{30} \right) = \frac{1}{3}$. The probability that your friend is between 10 and 20 minutes late is $\frac{1}{3}$.

Now Work Problem 13
Section 7.1 Properties of the Normal Distribution

We introduced the uniform density function so that we could associate probability with area. We are now better prepared to discuss the most frequently used continuous distribution, the normal distribution.

**Graph a Normal Curve**

Many continuous random variables, such as IQ scores, birth weights of babies, or weights of M&Ms, have relative frequency histograms that have a shape similar to Figure 3. Relative frequency histograms that have a shape similar to Figure 3 are said to have the shape of a normal curve.

A continuous random variable is **normally distributed**, or has a normal probability distribution, if the relative frequency histogram of the random variable has the shape of a normal curve.

**Definition**

A continuous random variable is normally distributed, or has a normal probability distribution, if the relative frequency histogram of the random variable has the shape of a normal curve.

**Historical Note**

Karl Pearson coined the phrase normal curve. He did not do this to imply that a distribution that is not normal is abnormal. Rather, Pearson wanted to avoid giving the name of the distribution a proper name, such as Gaussian (as in Carl Friedrich Gauss).

Figure 4 shows a normal curve, demonstrating the roles that \( \mu \) and \( \sigma \) play in drawing the curve. For any distribution, the mode represents the “high point” of the graph of the distribution. The median represents the point where 50% of the area under the distribution is to the left and 50% of the area under the distribution is to the right. The mean represents the balancing point of the graph of the distribution (see Figure 2 on page 131 in Section 3.1). For symmetric distributions with a single peak, such as the normal distribution, the mean = median = mode. Because of this, the mean, \( \mu \), is the high point of the graph of the distribution.

The points at \( x = \mu - \sigma \) and \( x = \mu + \sigma \) are the inflection points on the normal curve. The inflection points are the points on the curve where the curvature of the graph changes. To the left of \( x = \mu - \sigma \) and to the right of \( x = \mu + \sigma \), the curve is drawn upward \( (\cup) \). In between \( x = \mu - \sigma \) and \( x = \mu + \sigma \), the curve is drawn downward \( (\cap) \).

Figure 5 shows how changes in \( \mu \) and \( \sigma \) change the position or shape of a normal curve. In Figure 5(a), two normal density curves are drawn with the location of the inflection points labeled. One density curve has \( \mu = 0, \sigma = 1 \), and the other has \( \mu = 3, \sigma = 1 \). We can see that increasing the mean from 0 to 3 caused the graph to shift three units to the right but maintained its shape. In Figure 5(b), two normal density curves are drawn, again with the inflection points labeled. One density curve has \( \mu = 0, \sigma = 1 \), and the other has \( \mu = 0, \sigma = 2 \). We can see that increasing the standard deviation from 1 to 2 causes the graph to become flatter and more spread out but maintained its location of center.

*The vertical scale on the graph, which indicates density, is purposely omitted. The vertical scale, while important, will not play a role in any of the computations using this curve.*
Chapter 7  The Normal Probability Distribution

### State the Properties of the Normal Curve

The normal probability density function satisfies all the requirements that are necessary to have a probability distribution. We list the properties of the normal density curve next.

#### Properties of the Normal Density Curve

1. It is symmetric about its mean, \( \mu \).
2. Because mean \( \mu \) = median \( m \) = mode, there is a single peak and the highest point occurs at \( x = \mu \).
3. It has inflection points at \( \mu - \sigma \) and \( \mu + \sigma \).
4. The area under the curve is 1.
5. The area under the curve to the right of \( \mu \) equals the area under the curve to the left of \( \mu \), which equals \( \frac{1}{2} \).
6. As \( x \) increases without bound (gets larger and larger), the graph approaches, but never reaches, the horizontal axis. As \( x \) decreases without bound (gets larger and larger in the negative direction), the graph approaches, but never reaches, the horizontal axis.
7. The Empirical Rule: Approximately 68% of the area under the normal curve is between \( x = \mu - \sigma \) and \( x = \mu + \sigma \). Approximately 95% of the area under the normal curve is between \( x = \mu - 2\sigma \) and \( x = \mu + 2\sigma \). Approximately 99.7% of the area under the normal curve is between \( x = \mu - 3\sigma \) and \( x = \mu + 3\sigma \). See Figure 6.

### Explain the Role of Area in the Normal Density Function

Let’s look at an example of a normally distributed random variable.

#### Example 3  A Normal Random Variable

**Problem:** The relative frequency distribution given in Table 1 represents the heights of a pediatrician’s 200 three-year-old female patients. The raw data indicate that the mean height of the patients is \( \mu = 38.72 \) inches with standard deviation \( \sigma = 3.17 \) inches.

**a)** Draw a relative frequency histogram of the data. Comment on the shape of the distribution.
(b) Draw a normal curve with $\mu = 38.72$ inches and $\sigma = 3.17$ inches on the relative frequency histogram. Compare the area of the rectangle for heights between 40 and 40.9 inches to the area under the normal curve for heights between 40 and 40.9 inches.

**Approach**

(a) Draw the relative frequency histogram. If the histogram is shaped like Figure 3, we say that height is approximately normal. We say “approximately normal,” rather than “normal,” because the normal curve is an “idealized” description of the data, and data rarely follow the curve exactly.

(b) Draw the normal curve on the histogram with the high point at $\mu$ and the inflection points at $\mu - \sigma$ and $\mu + \sigma$. Shade the rectangle corresponding to heights between 40 and 40.9 inches, and compare the area of the shaded region to the area under the normal curve between 40 and 40.9.

**Solution**

(a) Figure 7 shows the relative frequency distribution. The relative frequency histogram is symmetric and bell-shaped.

(b) In Figure 8, the normal curve with $\mu = 38.72$ and $\sigma = 3.17$ is superimposed on the relative frequency histogram. The figure demonstrates that the normal curve describes the heights of 3-year-old females fairly well. We conclude that the heights of 3-year-old females are approximately normal with $\mu = 38.72$ and $\sigma = 3.17$.

Figure 8 also shows the rectangle corresponding to heights between 40 and 40.9 inches. The area of this rectangle represents the proportion of 3-year-old females with heights between 40 and 40.9 inches. Notice that the area of this shaded region is very close to the area under the normal curve for the same region, so we can use the area under the normal curve to approximate the proportion of 3-year-old females with heights between 40 and 40.9 inches!

**In Other Words**

Models are not always mathematical. For example, a map can be thought of as a model of a highway system. The model does not show every detail of the highway system (such as traffic lights), but it does serve the purpose of describing how to get from point A to point B. Mathematical models do the same thing: They make assumptions to simplify the mathematics, while still trying to accomplish the goal of accurately describing reality.

The normal curve drawn in Figure 8 is a model. In mathematics, a model is an equation, table, or graph that is used to describe reality. The normal distribution, or normal curve, is a model that is used to describe variables that are approximately normally distributed. For example, the normal curve drawn in Figure 8 does a good job of describing the observed distribution of heights of 3-year-old females.
The equation (model) that is used to determine the probability of a continuous random variable is called a probability density function (or pdf). The normal probability density function is given by

\[ y = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]

where \( \mu \) is the mean and \( \sigma \) is the standard deviation of the normal random variable. This equation represents the normal distribution. Don’t feel threatened by this equation, because we will not be using it in this text. Instead, we will use the normal distribution in graphical form by drawing the normal curve, as we did in Figure 5.

We now summarize the role area plays in the normal curve.

**Area under a Normal Curve**

Suppose that a random variable \( X \) is normally distributed with mean \( \mu \) and standard deviation \( \sigma \). The area under the normal curve for any interval of values of the random variable \( X \) represents either

- the proportion of the population with the characteristic described by the interval of values or
- the probability that a randomly selected individual from the population will have the characteristic described by the interval of values.

### EXAMPLE 4 Interpreting the Area under a Normal Curve

**Problem:** The serum total cholesterol for males 20 to 29 years old is approximately normally distributed with mean \( \mu = 180 \) and standard deviation \( \sigma = 36.2 \), based on data obtained from the National Health and Nutrition Examination Survey.

(a) Draw a normal curve with the parameters labeled.

(b) An individual with total cholesterol greater than 200 is considered to have high cholesterol. Shade the region under the normal curve to the right of \( x = 200 \).

(c) Suppose that the area under the normal curve to the right of \( x = 200 \) is 0.2903. (You will learn how to find this area in Section 7.3.) Provide two interpretations of this result.

**Approach**

(a) Draw the normal curve with the mean \( \mu = 180 \) labeled at the high point and the inflection points at \( \mu - \sigma = 180 - 36.2 = 143.8 \) and \( \mu + \sigma = 180 + 36.2 = 216.2 \).

(b) Shade the region under the normal curve to the right of \( x = 200 \).

(c) The two interpretations of this shaded region are (1) the proportion of 20- to 29-year-old males who have high cholesterol and (2) the probability that a randomly selected 20- to 29-year-old male has high cholesterol.

**Solution**

(a) Figure 9(a) shows the graph of the normal curve.
(b) Figure 9(b) shows the region under the normal curve to the right of $x = 200$ shaded.

(c) The two interpretations for the area of this shaded region are (1) the proportion of 20- to 29-year-old males that have high cholesterol is 0.2903 and (2) the probability that a randomly selected 20- to 29-year-old male has high cholesterol is 0.2903.

Describe the Relation between a Normal Random Variable and a Standard Normal Random Variable

At this point, we know that a random variable $X$ is approximately normally distributed if its relative frequency histogram has the shape of a normal curve. We use a normal random variable with mean $\mu$ and standard deviation $\sigma$ to model the distribution of $X$. The area below the normal curve (model of $X$) represents the proportion of the population with a given characteristic or the probability that a randomly selected individual from the population will have a given characteristic.

The question now becomes, “How do I find the area under the normal curve?” Finding the area under a curve requires techniques introduced in calculus, which are beyond the scope of this text. An alternative would be to use a series of tables to find areas. However, this would result in an infinite number of tables being created for each possible mean and standard deviation!

A solution to the problem lies in the $z$-score. Recall that the $z$-score allows us to transform a random variable $X$ with mean $\mu$ and standard deviation $\sigma$ into a random variable $Z$ with mean 0 and standard deviation 1.

Standardizing a Normal Random Variable

Suppose that the random variable $X$ is normally distributed with mean $\mu$ and standard deviation $\sigma$. Then the random variable

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

is normally distributed with mean $\mu = 0$ and standard deviation $\sigma = 1$. The random variable $Z$ is said to have the standard normal distribution.

This result is powerful! We need only one table of areas corresponding to the standard normal distribution. If a normal random variable has mean different from 0 or standard deviation different from 1, we transform the normal random variable into a standard normal random variable $Z$, and then we use a table to find the area and, therefore, the probability.

We demonstrate the idea behind standardizing a normal random variable in the next example.

EXAMPLE 5

Relation between a Normal Random Variable and a Standard Normal Random Variable

Problem: The heights of a pediatrician’s 200 three-year-old female patients are approximately normal with mean $\mu = 38.72$ inches and $\sigma = 3.17$ inches. We wish to demonstrate that the area under the normal curve between 35 and 38 inches is equal
Chapter 7  The Normal Probability Distribution

to the area under the standard normal curve between the z-scores corresponding to heights of 35 and 38 inches.

**Approach**

**Step 1:** Draw a normal curve and shade the area representing the proportion of 3-year-old females between 35 and 38 inches tall.

**Step 2:** Standardize the values \( x = 35 \) and \( x = 38 \) using

\[
 z = \frac{x - \mu}{\sigma}
\]

**Step 3:** Draw the standard normal curve with the standardized values of \( x = 35 \) and \( x = 38 \) labeled. Shade the area that represents the proportion of 3-year-old females between 35 and 38 inches tall. Comment on the relation between the two shaded regions.

**Solution**

**Step 1:** Figure 10(a) shows the normal curve with mean \( \mu = 38.72 \) and \( \sigma = 3.17 \). The region between \( x = 35 \) and \( x = 38 \) is shaded.

**Step 2:** With \( \mu = 38.72 \) and \( \sigma = 3.17 \), the standardized version of \( x = 35 \) is

\[
 z = \frac{x - \mu}{\sigma} = \frac{35 - 38.72}{3.17} = -1.17
\]

The standardized version of \( x = 38 \) is

\[
 z = \frac{x - \mu}{\sigma} = \frac{38 - 38.72}{3.17} = -0.23
\]

**Step 3:** Figure 10(b) shows the standard normal curve with the region between \( z = -1.17 \) and \( z = -0.23 \) shaded.

**Figure 10**

The area under the normal curve with \( \mu = 38.72 \) inches and \( \sigma = 3.17 \) inches bounded to the left by \( x = 35 \) and bounded to the right by \( x = 38 \) is equal to the area under the standard normal curve bounded to the left by \( z = -1.17 \) and bounded to the right by \( z = -0.23 \).
7.1 Assess Your Understanding

Concepts and Vocabulary

1. State the two characteristics of the graph of a probability density function.

2. To find the probabilities for continuous random variables, we do not use probability distribution functions, but instead we use probability density functions.

3. Provide two interpretations of the area under the graph of a probability density function.

4. Why do we standardize normal random variables to find the area under any normal curve?

5. The points at \( x = \frac{\mu - \sigma}{\sigma} \) and \( x = \frac{\mu + \sigma}{\sigma} \) are the inflection points on the normal curve.

6. As \( \sigma \) increases, the normal density curve becomes more spread out. Knowing that the area under the density curve must be 1, what effect does increasing \( \sigma \) have on the height of the curve?

For Problems 7–12, determine whether the graph can represent a normal density function. If it cannot, explain why.

7. No

8. Yes

9. No

10. No

11. Yes

12. No

Skill Building

Problems 13–16 use the information presented in Examples 1 and 2.

13. Find the probability that your friend is between 5 and 10 minutes late. \( \frac{1}{6} \)

14. Find the probability that your friend is between 15 and 25 minutes late. \( \frac{1}{3} \)

15. Find the probability that your friend is at least 20 minutes late. \( \frac{1}{3} \)

16. Find the probability that your friend is no more than 5 minutes late. \( \frac{1}{6} \)

17. Uniform Distribution The random-number generator on calculators randomly generates a number between 0 and 1. The random variable \( X \), the number generated, follows a uniform probability distribution.

(a) Draw the graph of the uniform density function.

(b) What is the probability of generating a number between 0 and 0.2? \( 0.2 \)

(c) What is the probability of generating a number between 0.25 and 0.6? \( 0.35 \)

(d) What is the probability of generating a number greater than 0.95? \( 0.05 \)

(e) Use your calculator or statistical software to randomly generate 200 numbers between 0 and 1. What proportion of the numbers are between 0 and 0.2? Compare the result with part (b).
18. **Uniform Distribution** The reaction time $X$ (in minutes) of a certain chemical process follows a uniform probability distribution with $5 \leq X \leq 10$.
   (a) Draw the graph of the density curve.
   (b) What is the probability that the reaction time is between 6 and 8 minutes? \[ \frac{2}{5} \]
   (c) What is the probability that the reaction time is between 5 and 8 minutes? \[ \frac{3}{5} \]
   (d) What is the probability that the reaction time is less than 6 minutes? \[ \frac{1}{5} \]

In Problems 19–22, determine whether or not the histogram indicates that a normal distribution could be used as a model for the variable.

19. **Birth Weights** The following relative frequency histogram represents the birth weights (in grams) of babies whose term was 36 weeks. Normal

![Birth Weights of Babies Whose Term Was 36 Weeks](image)

20. **Waiting in Line** The following relative frequency histogram represents the waiting times in line (in minutes) for the Demon Roller Coaster for 2000 randomly selected people on a Saturday afternoon in the summer. Not normal

![Waiting Time for the Demon Roller Coaster](image)

21. **Length of Phone Calls** The following relative frequency histogram represents the length of phone calls on my wife’s cell phone during the month of September. Not normal

![Length of Phone Calls](image)

22. **Incubation Times** The following relative frequency histogram represents the incubation times of a random sample of Rhode Island Red hens’ eggs. Normal

![Rhode Island Red Hen Incubation Times](image)

23. One graph in the following figure represents a normal distribution with mean $\mu = 10$ and standard deviation $\sigma = 3$. The other graph represents a normal distribution with mean $\mu = 10$ and standard deviation $\sigma = 2$. Determine which graph is which and explain how you know.
   A: $\mu = 10$, $\sigma = 3$; B: $\mu = 10$, $\sigma = 2$

![Normal Distribution Graphs](image)

24. One graph in the following figure represents a normal distribution with mean $\mu = 8$ and standard deviation $\sigma = 2$. The other graph represents a normal distribution with mean $\mu = 14$ and standard deviation $\sigma = 2$. Determine which graph is which and explain how you know.
   A: $\mu = 8$, $\sigma = 2$; B: $\mu = 14$, $\sigma = 2$
Section 7.1 Properties of the Normal Distribution

Applying the Concepts

29. You Explain It! Cell Phone Rates Monthly charges for cell phone plans in the United States are normally distributed with mean \( \mu = 62 \) and standard deviation \( \sigma = 18 \). 

   Source: Based on information from Consumer Reports

30. You Explain It! Refrigerators The lives of refrigerators are normally distributed with mean \( \mu = 14 \) years and standard deviation \( \sigma = 2.5 \) years.

   Source: Based on information from Consumer Reports

   (a) Draw a normal curve with the parameters labeled.
   (b) Shade the region that represents the proportion of refrigerators that last for more than 17 years.
   (c) Suppose that the area under the normal curve to the right of \( x = 17 \) is 0.1151. Provide two interpretations of this result.

31. You Explain It! Birth Weights The birth weights of full-term babies are normally distributed with mean \( \mu = 3,400 \) grams and \( \sigma = 505 \) grams.

   Source: Based on data obtained from the National Vital Statistics Report, Vol. 48, No. 3

   (a) Draw a normal curve with the parameters labeled.
   (b) Shade the region that represents the proportion of full-term babies who weigh more than 4,410 grams.
   (c) Suppose that the area under the normal curve to the right of \( x = 4,410 \) is 0.0228. Provide two interpretations of this result.

32. You Explain It! Height of 10-Year-Old Males The heights of 10-year-old males are normally distributed with mean \( \mu = 55.9 \) inches and \( \sigma = 5.7 \) inches.

   (a) Draw a normal curve with the parameters labeled.
   (b) Shade the region that represents the proportion of 10-year-old males who are less than 46.5 inches tall.
   (c) Suppose that the area under the normal curve to the left of \( x = 46.5 \) is 0.0496. Provide two interpretations of this result.

33. You Explain It! Gestation Period The lengths of human pregnancies are normally distributed with \( \mu = 266 \) days and \( \sigma = 16 \) days.

   (a) The following figure represents the normal curve with \( \mu = 266 \) days and \( \sigma = 16 \) days. The area to the right of \( x = 280 \) is 0.1908. Provide two interpretations of this area.
   (b) The following figure represents the normal curve with \( \mu = 266 \) days and \( \sigma = 16 \) days. The area between
384 Chapter 7 The Normal Probability Distribution

\[ x = 230 \text{ and } x = 260 \text{ is } 0.3416. \text{ Provide two interpretations of this area.} \]

![Area = 0.3416]

34. You Explain It! Miles per Gallon

Elena conducts an experiment in which she fills up the gas tank on her Toyota Camry 40 times and records the miles per gallon for each fill-up. A histogram of the miles per gallon indicates that a normal distribution with mean of 24.6 miles per gallon and a standard deviation of 3.2 miles per gallon could be used to model the gas mileage for her car.

(a) The following figure represents the normal curve with \( \mu = 24.6 \) miles per gallon and \( \sigma = 3.2 \) miles per gallon. The area under the curve to the right of \( x = 26 \) is 0.3309. Provide two interpretations of this area.

![Area = 0.3309]

(b) The following figure represents the normal curve with \( \mu = 24.6 \) miles per gallon and \( \sigma = 3.2 \) miles per gallon. The area under the curve between \( x = 18 \) and \( x = 21 \) is 0.1107. Provide two interpretations of this area.

![Area = 0.1107]

35. A random variable \( X \) is normally distributed with \( \mu = 10 \) and \( \sigma = 3 \).

(a) Compute \( z_1 = \frac{x_1 - \mu}{\sigma} \) for \( x_1 = 8 \). \(-0.67\)

(b) Compute \( z_2 = \frac{x_2 - \mu}{\sigma} \) for \( x_2 = 12 \). \(0.67\)

(c) The area under the normal curve between \( x_1 = 8 \) and \( x_2 = 12 \) is 0.495. What is the area between \( z_1 \) and \( z_2 \)?

36. A random variable \( X \) is normally distributed with \( \mu = 25 \) and \( \sigma = 6 \).

(a) Compute \( z_1 = \frac{x_1 - \mu}{\sigma} \) for \( x_1 = 18 \). \(-1.17\)

(b) Compute \( z_2 = \frac{x_2 - \mu}{\sigma} \) for \( x_2 = 30 \). \(0.83\)

(c) The area under the normal curve between \( x_1 = 18 \) and \( x_2 = 30 \) is 0.6760. What is the area between \( z_1 \) and \( z_2 \)?

37. Hitting a Pitching Wedge

In the game of golf, distance control is just as important as how far a player hits the ball. Michael went to the driving range with his range finder and hit 75 golf balls with his pitching wedge and measured the distance each ball traveled (in yards). He obtained the following data:

<table>
<thead>
<tr>
<th>Distance (yards)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 97 101 101 103 100 99 100 100</td>
</tr>
<tr>
<td>104 100 101 98 100 99 97 101</td>
</tr>
<tr>
<td>98 94 98 107 98 100 98 103 100</td>
</tr>
<tr>
<td>98 94 104 104 98 101 99 97 103</td>
</tr>
<tr>
<td>102 101 100 95 104 99 102 95</td>
</tr>
<tr>
<td>99 102 103 97 101 102 96 102 99</td>
</tr>
<tr>
<td>96 108 103 100 95 101 103 105 100</td>
</tr>
<tr>
<td>94 99 95</td>
</tr>
</tbody>
</table>

(a) Use Minitab or some other statistical software to construct a relative frequency histogram. Comment on the shape of the distribution.

(b) Use Minitab or some other statistical software to draw a normal density curve on the relative frequency histogram.

(c) Do you think the normal density function accurately describes the distance Michael hits with a pitching wedge? Why?

38. Heights of 5-Year-Old Females

The following frequency distribution represents the heights (in inches) of eighty randomly selected 5-year-old females.

<table>
<thead>
<tr>
<th>Height (inches)</th>
</tr>
</thead>
<tbody>
<tr>
<td>44.5 42.4 42.2 46.2 45.7 44.8 43.3 39.5</td>
</tr>
<tr>
<td>45.4 43.0 43.4 44.7 38.6 41.6 46.9</td>
</tr>
<tr>
<td>39.6 44.7 36.5 42.7 40.6 47.5 48.4 37.5</td>
</tr>
<tr>
<td>45.5 43.3 41.2 40.5 44.4 42.6 42.0 40.3</td>
</tr>
<tr>
<td>42.0 42.2 38.5 43.6 40.6 45.0 40.7 36.3</td>
</tr>
<tr>
<td>44.5 37.6 42.2 40.3 48.5 41.6 41.7 38.9</td>
</tr>
<tr>
<td>39.5 43.6 41.3 38.8 41.9 40.3 42.1 41.9</td>
</tr>
<tr>
<td>42.3 44.6 40.5 37.4 44.5 40.7 38.2 42.6</td>
</tr>
<tr>
<td>44.0 35.9 43.7 48.1 38.7 46.0 43.4 44.6</td>
</tr>
<tr>
<td>37.7 34.6 42.4 42.7 47.0 42.8 39.9 42.3</td>
</tr>
</tbody>
</table>

(a) Use Minitab or some other statistical software to construct a relative frequency histogram. Comment on the shape of the distribution.

(b) Use Minitab or some other statistical software to draw a normal density curve on the relative frequency histogram.

(c) Do you think the normal density function accurately describes the heights of 5-year-old females? Why?
In Section 7.1, we introduced the normal distribution. We learned that, if \( X \) is a normally distributed random variable, we can use the area under the normal density function to obtain the proportion of a population with a certain characteristic, or the probability that a randomly selected individual from the population has the characteristic. To find the area under the normal curve by hand, we first convert the random variable \( X \) to a standard normal random variable \( Z \) and find the area under the standard normal curve. In this section we discuss methods for finding area under the standard normal curve.

Properties of the Standard Normal Distribution

The standard normal distribution has a mean of 0 and a standard deviation of 1. The standard normal curve, therefore, has its high point located at 0 and inflection points located at \( \pm 1 \). We use \( Z \) to represent a standard normal random variable. The graph of the standard normal curve is presented in Figure 11.

Although we stated the properties of normal curves in Section 7.1, it is worthwhile to restate them here in terms of the standard normal curve.

Properties of the Standard Normal Curve

1. It is symmetric about its mean, \( \mu = 0 \), and has standard deviation \( \sigma = 1 \).
2. The mean = median = mode = 0. Its highest point occurs at \( z = 0 \).
3. It has inflection points at \( \mu - \sigma = 0 - 1 = -1 \) and \( \mu + \sigma = 0 + 1 = 1 \).
4. The area under the curve is 1.
5. The area under the curve to the right of \( \mu = 0 \) equals the area under the curve to the left of \( \mu = 0 \), which equals \( \frac{1}{2} \).
6. As the value of \( Z \) increases, the graph approaches, but never equals, zero. As the value of \( Z \) decreases, the graph approaches, but never equals, zero.
7. The Empirical Rule: Approximately 68\% of the area under the standard normal curve is between \( z = -1 \) and \( z = 1 \). Approximately 95\% of the area under the standard normal curve is between \( z = -2 \) and \( z = 2 \). Approximately 99.7\% of the area under the standard normal curve is between \( z = -3 \) and \( z = 3 \). See Figure 12.

We now discuss the procedure for finding area under the standard normal curve.

Find the Area under the Standard Normal Curve

We discuss two methods for finding area under the standard normal curve. The first method uses a table of areas that has been constructed for various values of \( Z \). The second method involves the use of statistical software or a graphing calculator with advanced statistical features.

Table V, which can be found on the inside back cover of the text and in Appendix A, gives areas under the standard normal curve for values to the left of a specified \( z \)-score, \( z \), as shown in Figure 13.
Chapter 7 The Normal Probability Distribution

The shaded region represents the area under the standard normal curve to the left of $z$. When finding area under a normal curve, you should always sketch the normal curve and shade the area in question. This practice will help to minimize errors.

**EXAMPLE 1**

**Finding Area under the Standard Normal Curve to the Left of a z-Score**

**Problem:** Find the area under the standard normal curve that lies to the left of $z = 1.68$.

**Approach**

**Step 1:** Draw a standard normal curve with $z = 1.68$ labeled, and shade the area under the curve to the left of $z = 1.68$.

**Step 2:** The rows in Table V represent the ones and tenths digits of $z$, while the columns represent the hundredths digit. To find the area under the curve to the left of $z = 1.68$, we split $1.68$ into $1.6$ and $0.08$. Find the row that represents $1.6$ and the column that represents $0.08$ in Table V. Identify where the row and column intersect. This value is the area.

**Solution**

**Step 1:** Figure 14 shows the graph of the standard normal curve with $z = 1.68$ labeled. The area to the left of $z = 1.68$ is shaded.

**Step 2:** A portion of Table V is presented in Figure 15. We have enclosed the row that represents $1.6$ and the column that represents $0.08$. The value located where the row and column intersect is the area we are seeking. The area to the left of $z = 1.68$ is $0.9535$.

The area under a standard normal curve may also be determined using statistical software or a graphing calculator with advanced statistical features.

**EXAMPLE 2**

**Finding the Area under a Standard Normal Curve Using Technology**

**Problem:** Find the area under the standard normal curve to the left of $z = 1.68$ using statistical software or a graphing calculator with advanced statistical features.

**Approach:** We will use MINITAB and a TI-84 Plus graphing calculator to find the area. The steps for determining the area under the standard normal curve using MINITAB, Excel, and the TI-83/84 Plus graphing calculators are given in the Technology Step-by-Step on page 396.

**Solution:** Figure 16(a) shows the results from MINITAB, and Figure 16(b) shows the results from a TI-84 Plus graphing calculator.
Notice that the output for MINITAB is titled Cumulative Distribution Function. Remember, the word cumulative means “less than or equal to,” so MINITAB is giving the area under the standard normal curve for values of $Z$ less than or equal to 1.68. The command required by the TI-84 Plus is `normalcdf(-1e99, 1.68, 0, 1)`. The cdf stands for cumulative distribution function. The TI-graphing calculators require a left and right bound. We use $1 \times 10^{99}$ for the left bound when obtaining areas “less than or equal to” some value.

![Figure 16](image)

**Cumulative Distribution Function**

Normal with mean $= 0$ and standard deviation $= 1$

$\begin{array}{l|l}
\mathbf{x} & P(\mathbf{X} \leq \mathbf{x}) \\
1.68 & 0.953521 \\
\end{array}$

(a)

(b)

In Other Words

Area right $= 1 -$ area left

**EXAMPLE 3**

Finding Area under the Standard Normal Curve to the Right of a $z$-Score

**Problem:** Find the area under the standard normal curve to the right of $z = -0.46$.

**Approach**

**Step 1:** Draw a standard normal curve with $z = -0.46$ labeled, and shade the area under the curve to the right of $z = -0.46$.

**Step 2:** Find the row that represents $-0.4$ and the column that represents 0.06 in Table V. Identify where the row and column intersect. This value is the area to the left of $z = -0.46$.

**Step 3:** The area under the standard normal curve to the right of $z = -0.46$ is 1 minus the area to the left of $z = -0.46$.

**Solution**

**Step 1:** Figure 17 shows the graph of the standard normal curve with $z = -0.46$ labeled. The area to the right of $z = -0.46$ is shaded.

**Step 2:** A portion of Table V is presented in Figure 18. We have enclosed the row that represents $-0.4$ and the column that represents 0.06. The value where the row and column intersect is the area to the left of $z = -0.46$. The area to the left of $z = -0.46$ is 0.3228.

**Figure 18**

<table>
<thead>
<tr>
<th>$z$</th>
<th>.00</th>
<th>.01</th>
<th>.02</th>
<th>.03</th>
<th>.04</th>
<th>.05</th>
<th>.06</th>
<th>.07</th>
<th>.08</th>
<th>.09</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0002</td>
</tr>
<tr>
<td>3.3</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0004</td>
<td>0.0004</td>
<td>0.0004</td>
<td>0.0004</td>
<td>0.0004</td>
<td>0.0004</td>
<td>0.0003</td>
</tr>
<tr>
<td>3.2</td>
<td>0.0007</td>
<td>0.0007</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0005</td>
<td>0.0005</td>
</tr>
<tr>
<td>3.1</td>
<td>0.0010</td>
<td>0.0009</td>
<td>0.0009</td>
<td>0.0009</td>
<td>0.0008</td>
<td>0.0008</td>
<td>0.0008</td>
<td>0.0008</td>
<td>0.0007</td>
<td>0.0007</td>
</tr>
<tr>
<td>3.0</td>
<td>0.0013</td>
<td>0.0013</td>
<td>0.0013</td>
<td>0.0012</td>
<td>0.0012</td>
<td>0.0011</td>
<td>0.0011</td>
<td>0.0011</td>
<td>0.0010</td>
<td>0.0010</td>
</tr>
<tr>
<td>0.5</td>
<td>0.3085</td>
<td>0.3050</td>
<td>0.3015</td>
<td>0.2981</td>
<td>0.2946</td>
<td>0.2912</td>
<td>0.2877</td>
<td>0.2843</td>
<td>0.2810</td>
<td>0.2776</td>
</tr>
<tr>
<td>0.4</td>
<td>0.3446</td>
<td>0.3409</td>
<td>0.3372</td>
<td>0.3336</td>
<td>0.3300</td>
<td>0.3264</td>
<td>0.3228</td>
<td>0.3192</td>
<td>0.3156</td>
<td>0.3121</td>
</tr>
<tr>
<td>0.3</td>
<td>0.3821</td>
<td>0.3783</td>
<td>0.3745</td>
<td>0.3707</td>
<td>0.3669</td>
<td>0.3632</td>
<td>0.3594</td>
<td>0.3557</td>
<td>0.3520</td>
<td>0.3483</td>
</tr>
<tr>
<td>0.2</td>
<td>0.4207</td>
<td>0.4168</td>
<td>0.4129</td>
<td>0.4090</td>
<td>0.4052</td>
<td>0.4013</td>
<td>0.3974</td>
<td>0.3936</td>
<td>0.3897</td>
<td>0.3859</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4602</td>
<td>0.4562</td>
<td>0.4522</td>
<td>0.4483</td>
<td>0.4443</td>
<td>0.4404</td>
<td>0.4364</td>
<td>0.4325</td>
<td>0.4286</td>
<td>0.4247</td>
</tr>
<tr>
<td>0.0</td>
<td>0.5000</td>
<td>0.4960</td>
<td>0.4920</td>
<td>0.4880</td>
<td>0.4840</td>
<td>0.4801</td>
<td>0.4761</td>
<td>0.4721</td>
<td>0.4681</td>
<td>0.4641</td>
</tr>
</tbody>
</table>
Step 3: The area under the standard normal curve to the right of $z = -0.46$ is 1 minus the area to the left of $z = -0.46$.

\[
\text{Area right of } z = -0.46 = 1 - (\text{area left of } z = -0.46) = 1 - 0.3228 = 0.6772
\]

The area to the right of $z = -0.46$ is 0.6772.

The next example presents a situation in which we are interested in the area between two $z$-scores.

**EXAMPLE 4**

**Find the Area under the Standard Normal Curve between Two $z$-Scores**

**Problem:** Find the area under the standard normal curve between $z = -1.35$ and $z = 2.01$.

**Approach**

**Step 1:** Draw a standard normal curve with $z = -1.35$ and $z = 2.01$ labeled. Shade the area under the curve between $z = -1.35$ and $z = 2.01$.

**Step 2:** Find the area to the left of $z = -1.35$. Find the area to the left of $z = 2.01$.

**Step 3:** The area under the standard normal curve between $z = -1.35$ and $z = 2.01$ is the area to the left of $z = 2.01$ minus the area to the left of $z = -1.35$.

**Solution**

**Step 1:** Figure 19 shows the standard normal curve with the area between $z = -1.35$ and $z = 2.01$ shaded.

**Step 2:** Based upon Table V, the area to the left of $z = -1.35$ is 0.0885. The area to the left of $z = 2.01$ is 0.9778.

**Step 3:** The area between $z = -1.35$ and $z = 2.01$ is

\[
\begin{align*}
\text{(Area between } z = -1.35 \text{ and } z = 2.01) &= (\text{area left of } z = 2.01) - (\text{area left of } z = -1.35) \\
&= 0.9778 - 0.0885 \\
&= 0.8893
\end{align*}
\]

The area between $z = -1.35$ and $z = 2.01$ is 0.8893.

We summarize the methods for obtaining area under the standard normal curve in Table 2.

Because the normal curve extends indefinitely in both directions on the $Z$-axis, there is no $z$-value for which the area under the curve to the left of the $z$-value is 1. For example, the area to the left of $z = 10$ is less than 1, even though graphing calculators and statistical software state that the area is 1, because they can only compute a limited number of decimal places. We will follow the practice of stating the area to the left of $z = -3.90$ or to the right of $z = 3.90$ as $<0.0001$. The area under the standard normal curve to the left of $z = 3.90$ or to the right of $z = -3.90$ will be stated as $>0.9999$. 

---

State the area under the standard normal curve to the left of $z = -3.90$ as $<0.0001$ (not 0). State the area under the standard normal curve to the left of $z = 3.90$ as $>0.9999$ (not 1).
Section 7.2  The Standard Normal Distribution

Table 2  Finding Areas under the Standard Normal Curve

<table>
<thead>
<tr>
<th>Problem</th>
<th>Approach</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find the area to the left of $z$.</td>
<td>Shade the area to the left of $z$.</td>
<td>Use Table V to find the row and column that correspond to $z$. The area is the value where the row and column intersect. Or, use technology to find the area.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Find the area to the right of $z$.</td>
<td>Shade the area to the right of $z$.</td>
<td>Use Table V to find the area left of $z$. The area to the right of $z$ is 1 minus the area to the left of $z$. Or, use technology to find the area.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Find the area between $z_1$ and $z_2$.</td>
<td>Shade the area between $z_1$ and $z_2$.</td>
<td>Use Table V to find the area to the left of $z_1$ and to the left of $z_2$. The area between $z_1$ and $z_2$ is $\left( \text{area to the left of } z_2 \right) - \left( \text{area to the left of } z_1 \right)$. Or, use technology to find the area.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Find $z$-Scores for a Given Area**

Up to this point, we have found areas given the value of a $z$-score. Often, we are interested in finding a $z$-score that corresponds to a given area. The procedure to follow is the reverse of the procedure for finding areas given $z$-scores.

**Example 5  Finding a $z$-Score from a Specified Area to Its Left**

**Problem:** Find the $z$-score so that the area to the left of the $z$-score is 0.32.

**Approach**

**Step 1:** Draw a standard normal curve with the area and corresponding unknown $z$-score labeled.

**Step 2:** Look for the area in the table closest to 0.32.

**Step 3:** Find the $z$-score that corresponds to the area closest to 0.32.

**Solution**

**Step 1:** Figure 20 shows the graph of the standard normal curve with the area of 0.32 labeled. We know that $z$ must be less than 0. Do you know why?

**Step 2:** We refer to Table V and look in the body of the table for an area closest to 0.32. The area closest to 0.32 is 0.3192. Figure 21 on the following page shows a portion of Table V with 0.3192 labeled.
Chapter 7 The Normal Probability Distribution

The $z$-score that corresponds to an area of 0.32 to its left is $z = -0.47$. See Figure 22.

Statistical software or graphing calculators with advanced statistical features can also be used to determine a $z$-score corresponding to a specified area.

Example 6 Finding a $z$-Score from a Specified Area to Its Left Using Technology

Problem: Find the $z$-score such that the area to the left of the $z$-score is 0.32 using statistical software or a graphing calculator with advanced statistical features.

Approach: We will use Excel and a TI-84 Plus graphing calculator to find the $z$-score. The steps for determining the $z$-score for MINITAB, Excel, and the TI-83/84 Plus graphing calculators are given in the Technology Step-by-Step on page 396.

Solution: Figure 23(a) shows the results from Excel, and Figure 23(b) shows the results from a TI-84 Plus graphing calculator.

The $z$-score that corresponds to an area of 0.32 to its left is $-0.47$, so $z = -0.47$.

It is useful to remember that if the area to the left of the $z$-score is less than 0.5 the $z$-score must be less than 0. If the area to the left of the $z$-score is greater than 0.5, the $z$-score must be greater than 0.
The next example deals with situations in which the area to the right of some unknown $z$-score is given. The solution uses the fact that the total area under the normal curve is 1.

**EXAMPLE 7**

**Finding a $z$-Score from a Specified Area to Its Right**

**Problem:** Find the $z$-score so that the area to the right of the $z$-score is 0.4332.

**Approach**

**Step 1:** Draw a standard normal curve with the area and corresponding unknown $z$-score labeled.

**Step 2:** Determine the area to the left of the unknown $z$-score.

**Step 3:** Look for the area in the table closest to the area determined in Step 2 and record the $z$-score that corresponds to the closest area.

**Solution**

**Step 1:** Figure 24 shows the standard normal curve with the area and unknown $z$-score labeled.

**Step 2:** Since the area under the entire normal curve is 1, the area to the left of the unknown $z$-score is $1$ minus the area right of the unknown $z$-score. Therefore,

$$\text{Area to the left} = 1 - \text{area to the right} = 1 - 0.4332 = 0.5668$$

**Step 3:** We look in the body of Table V for an area closest to 0.5668. See Figure 25. The area closest to 0.5668 is 0.5675.

The approximate $z$-score that corresponds to a right-tail area of 0.4332 is 0.17. Therefore, $z = 0.17$. See Figure 26.

**Now Work Problem 19**

In upcoming chapters, we will be interested in finding $z$-scores that separate the middle area of the standard normal curve from the area in its tails.

**EXAMPLE 8**

**Finding $z$-Scores from an Area in the Middle**

**Problem:** Find the $z$-scores that separates the middle 90% of the area under the standard normal curve from the area in the tails.

**Approach**

**Step 1:** Draw a standard normal curve with the middle 90% = 0.9 of the area separated from the area of $5\% = 0.05$ in each of the two tails. Label the unknown $z$-scores $z_1$ and $z_2$. 
Step 2: Look in the body of Table V to find the area closest to 0.05.

Step 3: Determine the \( z \)-score for the left tail.

Step 4: The area to the right of \( z_2 \) is 0.05. Therefore, the area to the left of \( z_2 \) is 0.95. Look in Table V for an area of 0.95 and find the corresponding \( z \)-score.

Solution

Step 1: Figure 27 shows the standard normal curve with the middle 90% of the area separated from the area in the two tails.

Step 2: We look in the body of Table V for an area closest to 0.05. See Figure 28. Notice that 0.0495 and 0.0505 are equally close to 0.05.

Step 3: We agree to take the mean of the two \( z \)-scores corresponding to the areas. The \( z \)-score corresponding to an area of 0.0495 is \(-1.65\). The \( z \)-score corresponding to an area of 0.0505 is \(-1.64\). Therefore, the approximate \( z \)-score corresponding to an area of 0.05 to its left is

\[
z = \frac{-1.65 + (-1.64)}{2} = -1.645
\]

Step 4: The area to the right of \( z_2 \) is 0.05. Therefore, the area to the left of \( z_2 = 1 - 0.05 = 0.95 \). In Table V, we find an area of 0.9495 corresponding to \( z = 1.64 \) and an area of 0.9505 corresponding to \( z = 1.65 \). Consequently, the approximate \( z \)-score corresponding to an area of 0.05 to the right is

\[
z_2 = \frac{1.65 + 1.64}{2} = 1.645
\]

See Figure 29.

We could also obtain the solution to Example 8 using symmetry. Because the standard normal curve is symmetric about its mean, 0, the \( z \)-score that corresponds to an area of 0.05 to the left will be the additive inverse (i.e., opposite) of the \( z \)-score that corresponds to an area of 0.05 to the right. Since the area to the left of \( z = -1.65 \) is 0.05, the area to the right of \( z = 1.65 \) is also 0.05.

We are often interested in finding the \( z \)-score that has a specified area to the right. For this reason, we have special notation to represent this situation.

The notation \( z_\alpha \) (pronounced “\( z \) sub alpha”) is the \( z \)-score such that the area under the standard normal curve to the right of \( z_\alpha \) is \( \alpha \). Figure 30 illustrates the notation.
EXAMPLE 9

**Finding the Value of \( z_{0.10} \)**

**Problem:** Find the value of \( z_{0.10} \).

**Approach:** We wish to find the \( z \)-value such that the area under the standard normal curve to the right of the \( z \)-value is 0.10.

**Technology Solution:** The area to the right of the unknown \( z \)-value is 0.10, so the area to the left is \( 1 - 0.10 = 0.90 \). A TI-84 Plus is used to find that the \( z \)-value such that the area to the left is 0.90 is 1.28. See Figure 31. Therefore, \( z_{0.10} = 1.28 \).

Figure 32 shows the \( z \)-value on the normal curve.

**Example 10**

**Finding Probabilities of Standard Normal Random Variables**

**Problem:** Evaluate: \( P(Z < 1.26) \)

**Approach**

**Step 1:** Draw a standard normal curve with the area that we desire shaded.
Step 2: Use Table V to find the area of the shaded region. This area represents the probability.

Solution

Step 1: Figure 33 shows the standard normal curve with the area to the left of $z = 1.26$ shaded.

Step 2: Using Table V, we find that the area under the standard normal curve to the left of $z = 1.26$ is 0.8962. Therefore, $P(Z < 1.26) = 0.8962$.

For any continuous random variable, the probability of observing a specific value of the random variable is 0. For example, for a standard normal random variable, $P(a) = 0$ for any value of $a$. This is because there is no area under the standard normal curve associated with a single value, so the probability must be 0. Therefore, the following probabilities are equivalent:

$$P(a < Z < b) = P(a \leq Z < b) = P(a < Z \leq b) = P(a \leq Z \leq b)$$

For example, $P(Z < 1.26) = P(Z \leq 1.26) = 0.8962$.

### 7.2 ASSESS YOUR UNDERSTANDING

**Concepts and Vocabulary**

1. State the properties of the standard normal curve.
2. If the area under the standard normal curve to the left of $z = 1.20$ is 0.8849, what is the area under the standard normal curve to the right of $z = 1.20$? **0.1151**
3. **True or False**: The area under the standard normal curve to the left of $z = 5.30$ is 1. Support your answer. False
4. Explain why $P(Z < -1.30) = P(Z \leq -1.30)$.

**Skill Building**

*In Problems 5–12, find the indicated areas. For each problem, be sure to draw a standard normal curve and shade the area that is to be found.*

5. Determine the area under the standard normal curve that lies to the left of
   - (a) $z = -2.45$ **0.0071**
   - (b) $z = -0.43$ **0.3336**
   - (c) $z = 1.35$ **0.9115**
   - (d) $z = 3.49$ **0.9995**

6. Determine the area under the standard normal curve that lies to the left of
   - (a) $z = -3.49$ **0.0002**
   - (b) $z = -1.99$ **0.0233**
   - (c) $z = 0.92$ **0.8212**
   - (d) $z = 2.90$ **0.9991**

7. Determine the area under the standard normal curve that lies to the right of
   - (a) $z = -3.01$ **0.9987**
   - (b) $z = -1.59$ **0.8441**

8. Determine the area under the standard normal curve that lies to the right of
   - (a) $z = -3.49$ **0.9999**
   - (b) $z = -0.55$ **0.7088**
   - (c) $z = 2.23$ **0.0129**
   - (d) $z = 3.45$ **0.0003**

9. Determine the area under the standard normal curve that lies between
   - (a) $z = -2.04$ and $z = 2.04$ **0.9586**
   - (b) $z = -0.55$ and $z = 0$ **0.2088**
   - (c) $z = -1.04$ and $z = 2.76$ **0.8479**

10. Determine the area under the standard normal curve that lies between
    - (a) $z = -2.55$ and $z = 2.55$ **0.9892**
    - (b) $z = -1.67$ and $z = 0$ **0.4525**
    - (c) $z = -3.03$ and $z = 1.98$ **0.9749**

11. Determine the total area under the standard normal curve
    - (a) to the left of $z = -2$ or to the right of $z = 2$ **0.0456**
    - (b) to the left of $z = -1.56$ or to the right of $z = 2.56$ **0.2028**
    - (c) to the left of $z = -0.24$ or to the right of $z = 1.20$ **0.2028**

12. Determine the total area under the standard normal curve
    - (a) to the left of $z = -2.94$ or to the right of $z = 2.94$ **0.0456**
    - (b) to the left of $z = -1.68$ or to the right of $z = 3.05$ **0.2028**
    - (c) to the left of $z = -0.88$ or to the right of $z = 1.23$ **0.2028**

   11. (b) **0.0646**  11. (c) **0.5203**  12. (a) **0.0032**
   12. (b) **0.0476**  12. (c) **0.2987**
In Problems 13 and 14, find the area of the shaded region for each standard normal curve.

13. (a) 0.8877

(b) 0.0500

(c) 0.9802

14. (a) 0.0198

(b) 0.3686

(c) 0.0234

In Problems 15–26, find the indicated z-score. Be sure to draw a standard normal curve that depicts the solution.

15. Find the z-score such that the area under the standard normal curve to its left is 0.1. \(-1.28\)

16. Find the z-score such that the area under the standard normal curve to its left is 0.2. \(-0.84\)

17. Find the z-score such that the area under the standard normal curve to its left is 0.98. \(2.05\)

18. Find the z-score such that the area under the standard normal curve to its left is 0.85. \(1.04\)

19. Find the z-score such that the area under the standard normal curve to its right is 0.25. \(0.67\)

20. Find the z-score such that the area under the standard normal curve to its right is 0.35. \(0.39\)

21. Find the z-score such that the area under the standard normal curve to its right is 0.89. \(-1.23\)

22. Find the z-score such that the area under the standard normal curve to its right is 0.75. \(-0.67\)

23. Find the z-scores that separate the middle 80% of the distribution from the area in the tails of the standard normal distribution. \(-1.28; 1.28\)

24. Find the z-scores that separate the middle 70% of the distribution from the area in the tails of the standard normal distribution. \(-1.04; 1.04\)

25. Find the z-scores that separate the middle 99% of the distribution from the area in the tails of the standard normal distribution. \(-2.575; 2.575\)

26. Find the z-scores that separate the middle 94% of the distribution from the area in the tails of the standard normal distribution. \(-1.64; 1.64\)

In Problems 27–32, find the value of \(z_a\).

27. \(z_{0.05} = 1.645\)

28. \(z_{0.35} = 0.39\)

29. \(z_{0.01} = 2.33\)

30. \(z_{0.02} = 2.05\)

31. \(z_{0.20} = 0.84\)

32. \(z_{0.15} = 1.04\)

In Problems 33–44, find the indicated probability of the standard normal random variable \(Z\).

33. \(P(Z < 1.93) = 0.9732\)

34. \(P(Z < -0.61) = 0.2709\)

35. \(P(Z > -2.98) = 0.9988\)

36. \(P(Z > 0.92) = 0.1766\)

37. \(P(-1.20 < Z < 2.34) = 0.8753\)

38. \(P(1.23 < Z < 1.56) = 0.0499\)

39. \(P(Z \geq 1.84) = 0.0329\)

40. \(P(Z \leq -0.92) = 0.8212\)

41. \(P(Z \leq 0.72) = 0.7642\)

42. \(P(Z < -2.69) = 0.0036\)
43. \( P(Z < -2.56 \text{ or } Z > 1.39) = 0.0875 \)
44. \( P(Z < -0.38 \text{ or } Z > 1.93) = 0.3788 \)

**Applying the Concepts**

45. **The Empirical Rule** The Empirical Rule states that about 68% of the data in a bell-shaped distribution lies within 1 standard deviation of the mean. For the standard normal distribution, this means about 68% of the data lies between \( z = -1 \) and \( z = 1 \). Verify this result. Verify that about 95% of the data lies within 2 standard deviations of the mean. Finally, verify that about 99.7% of the data lies within 3 standard deviations of the mean.

46. According to Table V, the area under the standard normal curve to the left of \( z = -1.34 \) is 0.0901. Without consulting Table V, determine the area under the standard normal curve to the right of \( z = 1.34 \). 0.0901

---

**TECHNOLOGY STEP-BY-STEP**

**The Standard Normal Distribution**

1. **TI-83/84 Plus**
   - **Finding Areas under the Standard Normal Curve**
     1. From the HOME screen, press \( 2^{\text{nd}} \) VARS to access the DISTRIBUTION menu.
     2. Select 2:normalcdf(.
     3. With normalcdf( on the HOME screen, type lowerbound, upperbound, 0, 1). For example, to find the area left of \( z = 1.26 \) under the standard normal curve, type

\[
\text{normalcdf}(-1E99, 1.26, 0, 1)
\]

and hit ENTER.

**Note:** When there is no lowerbound, enter \(-1E99\). When there is no upperbound, enter \(1E99\). The \( E \) shown is scientific notation; it is selected by pressing \( 2^{\text{nd}} \) then \( \times \).

2. **Finding Z-Scores Corresponding to an Area**
   - **From the HOME screen, press \( 2^{\text{nd}} \) VARS to access the DISTRIBUTION menu.**
   - **Select 3:invNorm(.**
   - **With invNorm( on the HOME screen, type “area left”, 0, 1).** For example, to find the \( z \)-score such that the area under the normal curve to the left of the \( z \)-score is 0.79, type

\[
\text{invNorm}(0.79, 0, 1)
\]

and hit ENTER.

**MINITAB**

1. **Finding Areas under the Standard Normal Curve**
   - **MINITAB** will find an area to the left of a specified \( z \)-score. Select the **Calc** menu, highlight **Probability Distributions**, and highlight **Normal. . .**

2. **Finding Z-Scores Corresponding to an Area**
   - **From the HOME screen,** select **NCDF**.
   - **With invNorm( on the HOME screen, type “area left”, 0, 1).** For example, to find the \( z \)-score such that the area under the normal curve to the left of the \( z \)-score is 0.79, type

\[
\text{invNorm}(0.79, 0, 1)
\]

and hit ENTER.

**Excel**

1. **Finding Areas under the Standard Normal Curve**
   - **Select Cumulative Probability.** Set the mean to 0 and the standard deviation to 1. Select **Input Constant**, and enter the specified \( z \)-score. Click OK.

2. **Finding Z-Scores Corresponding to an Area**
   - **Select Inverse Cumulative Probability.** Set the mean to 0 and the standard deviation to 1. Select **Input Constant**, and enter the specified area. Click OK.


Section 7.3 Applications of the Normal Distribution

Preparing for This Section Before getting started, review the following:

• Percentiles (Section 3.4, p. 168)

Objectives

1. Find and interpret the area under a normal curve
2. Find the value of a normal random variable

Find and Interpret the Area under a Normal Curve

From the discussions in Section 7.1, we know that finding the area under a normal curve by hand requires that we transform values of a normal random variable $X$ with mean $\mu$ and standard deviation $\sigma$ into values of the standard normal random variable $Z$ with mean 0 and standard deviation 1. This is accomplished by letting $Z = \frac{X - \mu}{\sigma}$ and using Table V to find the area under the standard normal curve. This idea is illustrated in Figure 34.

Now that we have the ability to find the area under a standard normal curve, we can find the area under any normal curve. We summarize the procedure next.

Finding the Area under Any Normal Curve

Step 1: Draw a normal curve and shade the desired area.
Step 2: Convert the values of $X$ to $z$-scores using $z = \frac{X - \mu}{\sigma}$.
Step 3: Draw a standard normal curve and shade the desired area.
Step 4: Find the area under the standard normal curve. This area is equal to the area under the normal curve drawn in Step 1.

Example 1 Finding Area under a Normal Curve

Problem: A pediatrician obtains the heights of her 200 three-year-old female patients. The heights are approximately normally distributed, with mean 38.72 inches and standard deviation 3.17 inches. Use the normal model to determine the proportion of the 3-year-old females that have a height less than 35 inches.

Approach: Follow Steps 1 through 4.
Chapter 7  The Normal Probability Distribution

Solution
Step 1: Figure 35 shows the normal curve with the area to the left of 35 shaded.
Step 2: We convert \( x = 35 \) to a \( z \)-score.
\[
z = \frac{x - \mu}{\sigma} = \frac{35 - 38.72}{3.17} = -1.17
\]
Step 3: Figure 36 shows the standard normal curve with the area to the left of \( z = -1.17 \) shaded. The area to the left of \( z = -1.17 \) is equal to the area to the left of \( x = 35 \).

Figure 35

Figure 36

Step 4: Using Table V, we find that the area to the left of \( z = -1.17 \) is 0.1210. The normal model indicates that the proportion of the pediatrician’s 3-year-old females that is less than 35 inches tall is 0.1210.

According to the results of Example 1, the proportion of 3-year-old females who are shorter than 35 inches is 0.1210. If the normal curve is a good model for determining proportions (or probabilities), then about 12.1% of the 200 three-year-olds in Table 1 should be shorter than 35 inches. For convenience, the information provided in Table 1 is repeated in Table 3.

From the relative frequency distribution in Table 3, we determine that 0.005 + 0.005 + 0.005 + 0.025 + 0.02 + 0.055 = 0.115 = 11.5% of the 3-year-old females are less than 35 inches tall. The results based on the normal curve are in close agreement with the actual results. The normal curve accurately models the heights of 3-year-old females.

Because the area under the normal curve represents a proportion, we can also use the area to find percentile ranks of scores. Recall that the \( k \)th percentile divides the lower \( k \)\% of a data set from the upper \((100 - k)\)\%.

Statistical software and graphing calculators with advanced statistical features can also be used to find areas under any normal curve.

Table 3

<table>
<thead>
<tr>
<th>Height (inches)</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>29.0–29.9</td>
<td>0.005</td>
</tr>
<tr>
<td>30.0–30.9</td>
<td>0.005</td>
</tr>
<tr>
<td>31.0–31.9</td>
<td>0.005</td>
</tr>
<tr>
<td>32.0–32.9</td>
<td>0.025</td>
</tr>
<tr>
<td>33.0–33.9</td>
<td>0.02</td>
</tr>
<tr>
<td>34.0–34.9</td>
<td>0.055</td>
</tr>
<tr>
<td>35.0–35.9</td>
<td>0.075</td>
</tr>
<tr>
<td>36.0–36.9</td>
<td>0.09</td>
</tr>
<tr>
<td>37.0–37.9</td>
<td>0.115</td>
</tr>
<tr>
<td>38.0–38.9</td>
<td>0.15</td>
</tr>
<tr>
<td>39.0–39.9</td>
<td>0.12</td>
</tr>
<tr>
<td>40.0–40.9</td>
<td>0.11</td>
</tr>
<tr>
<td>41.0–41.9</td>
<td>0.07</td>
</tr>
<tr>
<td>42.0–42.9</td>
<td>0.06</td>
</tr>
<tr>
<td>43.0–43.9</td>
<td>0.035</td>
</tr>
<tr>
<td>44.0–44.9</td>
<td>0.025</td>
</tr>
<tr>
<td>45.0–45.9</td>
<td>0.025</td>
</tr>
<tr>
<td>46.0–46.9</td>
<td>0.005</td>
</tr>
<tr>
<td>47.0–47.9</td>
<td>0.005</td>
</tr>
</tbody>
</table>

EXAMPLE 2
Finding Area under a Normal Curve Using Technology

Problem: Find the percentile rank of a 3-year-old female whose height is 43 inches using statistical software or a graphing calculator with advanced statistical features. From Example 1, we know that the heights are approximately normally distributed, with a mean of 38.72 inches and standard deviation of 3.17 inches.

Approach: We will use a TI-84 Plus graphing calculator to find the area. The steps for determining the area under any normal curve for MINITAB, Excel, and the
TI-83/84 Plus graphing calculators are given in the Technology Step-by-Step on pages 404–405.

Solution: Figure 37 shows the results from a TI-84 Plus graphing calculator. The area under the normal curve to the left of 43 is approximately 0.91. Therefore, 91% of the heights are less than 43 inches, and 9% of the heights are more than 43 inches. A 3-year-old female whose height is 43 inches is at the 91st percentile.

In Other Words
The normal probability density function is used to model random variables that appear to be normal (such as girls’ heights). A good model is one that yields results that are close to reality.

According to the relative frequency distribution in Table 3, the proportion of the 200 three-year-old females with heights between 35 and 40 inches is $0.075 + 0.09 + 0.115 + 0.15 + 0.12 = 0.55 = 55\%$. This is very close to the probability obtained in Example 3!
Find the Value of a Normal Random Variable

Often, rather than being interested in the proportion, probability or percentile for a value of a normal random variable, we are interested in calculating the value of a normal random variable required for the value to correspond to a certain proportion, probability, or percentile. For example, we might want to know the height of a 3-year-old girl at the 20th percentile. This means that we want to know the height of a 3-year-old girl who is taller than 20% of all 3-year-old girls.

**Procedure for Finding the Value of a Normal Random Variable Corresponding to a Specified Proportion, Probability, or Percentile**

**Step 1:** Draw a normal curve and shade the area corresponding to the proportion, probability, or percentile.

**Step 2:** Use Table V to find the $z$-score that corresponds to the shaded area.

**Step 3:** Obtain the normal value from the formula $x = \mu + z\sigma$.

---

**EXAMPLE 4**

**Finding the Value of a Normal Random Variable**

**Problem:** The heights of a pediatrician’s 200 three-year-old females are approximately normally distributed, with mean 38.72 inches and standard deviation 3.17 inches. Find the height of a 3-year-old female at the 20th percentile. That is, find the height of a 3-year-old female that separates the bottom 20% from the top 80%.

**Approach:** We follow Steps 1 through 3.

**Solution**

**Step 1:** Figure 40 shows the normal curve with the unknown value of $X$ separating the bottom 20% of the distribution from the top 80% of the distribution.

**Step 2:** From Table V, the area closest to 0.20 is 0.2005. The corresponding $z$-score is $-0.84$.

**Step 3:** The height of a 3-year-old female that separates the bottom 20% of the data from the top 80% is computed as follows:

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

$$= 38.72 + (-0.84)(3.17)$$

$$= 36.1 \text{ inches}$$

The height of a 3-year-old female at the 20th percentile is 36.1 inches.

---

**EXAMPLE 5**

**Finding the Value of a Normal Random Variable Using Technology**

**Problem:** Use statistical software or a graphing calculator with advanced statistical features to verify the results of Example 4. That is, find the height for a 3-year-old female that is at the 20th percentile, assuming their heights are approximately normally distributed, with a mean of 38.72 inches and a standard deviation of 3.17 inches.

*The formula provided in Step 3 is simply the formula for computing a $z$-score, solved for $x$.

$$z = \frac{x - \mu}{\sigma}$$  
*Formula for standardizing a value, $x$, for a random variable $X$*

$$z\sigma = x - \mu$$  
*Multiply both sides by $\sigma$.*

$$x = \mu + z\sigma$$  
*Add $\mu$ to both sides.*
Section 7.3 Applications of the Normal Distribution

Approach: We will use MINITAB to find the height at the 20th percentile. The steps for determining the value of a normal random variable given an area for MINITAB, Excel, and the TI-83/84 Plus graphing calculators are given in the Technology Step-by-Step on pages 404–405.

Solution: Figure 41 shows the results obtained from MINITAB. The height of a 3-year-old female at the 20th percentile is 36.1 inches.

Figure 41
Inverse Cumulative Distribution Function
Normal with mean = 38.7200 and standard deviation = 3.17000

Solution
x
0.2000
36.0521

Now Work Problem 25(a)

EXAMPLE 6 Finding the Value of a Normal Random Variable

Problem: The heights of a pediatrician’s 200 three-year-old females are approximately normally distributed, with mean 38.72 inches and standard deviation 3.17 inches. The pediatrician wishes to determine the heights that separate the middle 98% of the distribution from the bottom 1% and top 1%. In other words, find the 1st and 99th percentiles.

Approach: We follow Steps 1 through 3 given on page 400.

Solution

Step 1: Figure 42 shows the normal curve with the unknown values of \( x \) separating the bottom and top 1% of the distribution from the middle 98% of the distribution.

Step 2: First, we will find the \( z \)-score that corresponds to an area of 0.01 to the left. From Table V, the area closest to 0.01 is 0.0099. The corresponding \( z \)-score is \( -2.33 \). The \( z \)-score that corresponds to an area of 0.01 to the right is the \( z \)-score that has area 0.99 to the left. The area closest to 0.99 is 0.9901. The corresponding \( z \)-score is 2.33.

Step 3: The height of a 3-year-old female that separates the bottom 1% of the distribution from the top 99% is

\[
x_1 = \mu + z\sigma
= 38.72 + (-2.33)(3.17)
= 31.3
\]

The height of a 3-year-old female that separates the top 1% of the distribution from the bottom 99% is

\[
x_2 = \mu + z\sigma
= 38.72 + (2.33)(3.17)
= 46.1
\]

A 3-year-old female whose height is less than 31.3 inches is in the bottom 1% of all 3-year-old females, and a 3-year-old female whose height is more than 46.1 inches is in the top 1% of all 3-year-old females. The pediatrician might use this information to identify those patients who have unusual heights.

Figure 42

CAUTION If you are given a value of the random variable and asked to find the probability, proportion, or percentile corresponding to the value, convert the value to a \( z \)-score using \( z = \frac{x - \mu}{\sigma} \) and find the area from the table. If you are using a TI-83/84 Plus graphing calculator, use normalcdf[.

If you are asked to find the value of a random variable corresponding to a probability, proportion, or percentile, find the area that represents the given probability, proportion, or percentile and use \( x = \mu + z\sigma \) to find the value of the random variable. If you are using a TI-83/84 Plus graphing calculator, use invNorm[.

Now Work Problem 25(b)
7.3 Assess Your Understanding

Concepts and Vocabulary
1. Describe the procedure for finding the area under any normal curve.
2. Describe the procedure for finding the value of a normal random variable corresponding to a probability, proportion, or percentile.

Skill Building

In Problems 3–12, assume that the random variable $X$ is normally distributed, with mean $\mu = 50$ and standard deviation $\sigma = 7$. Compute the following probabilities. Be sure to draw a normal curve with the area corresponding to the probability shaded.

3. $P(X > 35)$
4. $P(X > 65)$
5. $P(X \leq 45)$
6. $P(X \geq 58)$
7. $P(40 < X < 65)$
8. $P(56 < X < 68)$
9. $P(55 \leq X \leq 70)$
10. $P(40 \leq X \leq 49)$
11. $P(38 < X \leq 55)$
12. $P(56 \leq X < 66)$

In Problems 13–16, assume that the random variable $X$ is approximately normally distributed with a standard deviation of 1 day. Find each indicated percentile for $X$.

13. The 9th percentile
14. The 90th percentile
15. The 81st percentile
16. The 38th percentile

Applying the Concepts

17. Egg Incubation Times The mean incubation time of fertilized chicken eggs kept at 100.5°F in a still-air incubator is 21 days. Suppose that the incubation times are approximately normally distributed with a standard deviation of 1 day. 

Source: University of Illinois Extension

(a) What is the probability that a randomly selected fertilized chicken egg hatches in less than 20 days? $0.8587$
(b) What is the probability that a randomly selected fertilized chicken egg takes over 22 days to hatch? $0.1587$
(c) What is the probability that a randomly selected fertilized chicken egg hatches between 19 and 21 days? $0.4772$
(d) Would it be unusual for an egg to hatch in less than 18 days? Yes

18. Reading Rates The reading speed of sixth-grade students is approximately normal, with a mean speed of 125 words per minute and a standard deviation of 24 words per minute.

(a) What is the probability that a randomly selected sixth-grade student reads less than 100 words per minute? $0.1492$
(b) What is the probability that a randomly selected sixth-grade student reads more than 140 words per minute? $0.3189$
(c) What is the probability that a randomly selected sixth-grade student reads between 110 and 130 words per minute? $0.3679$
(d) Would it be unusual for a sixth-grader to read more than 200 words per minute? Yes

19. Chips Ahoy! Cookies The number of chocolate chips in an 18-ounce bag of Chips Ahoy! chocolate chip cookies is approximately normally distributed with a mean of 1,262 chips and standard deviation 118 chips according to a study by cadets of the U.S. Air Force Academy. 

Source: Brad Warner and Jim Rutledge, Chance 12(1): 10–14, 1999

(a) What is the probability that a randomly selected 18-ounce bag of Chips Ahoy! contains between 1,000 and 1,400 chocolate chips, inclusive? $0.8658$
(b) What is the probability that a randomly selected 18-ounce bag of Chips Ahoy! contains fewer than 1,000 chocolate chips? $0.0132$
(c) What proportion of 18-ounce bags of Chips Ahoy! contains more than 1,200 chocolate chips? $0.7019$
(d) What proportion of 18-ounce bags of Chips Ahoy! contains fewer than 1,125 chocolate chips? $0.1230$
(e) What is the percentile rank of an 18-ounce bag of Chips Ahoy! that contains 1,475 chocolate chips? (f) What is the percentile rank of an 18-ounce bag of Chips Ahoy! that contains 1,050 chocolate chips?

20. Wendy’s Drive-Through Fast-food restaurants spend quite a bit of time studying the amount of time cars spend in their drive-throughs. Certainly, the faster the cars get through the drive-through, the more the opportunity for making money. In 2007, QSR Magazine studied drive-through times for fast-food restaurants and Wendy’s had the best time, with a mean time spent in the drive-through of 138.5 seconds. Assuming drive-through times are normally distributed with a standard deviation of 29 seconds, answer the following.

(a) What is the probability that a randomly selected car will get through Wendy’s drive-through in less than 100 seconds? $0.0918$
(b) What is the probability that a randomly selected car will spend more than 160 seconds in Wendy’s drive-through? $0.6625$
(c) What proportion of cars spend between 2 and 3 minutes in Wendy’s drive-through? $0.8658$
(d) Would it be unusual for a car to spend more than 3 minutes in Wendy’s drive-through? Why? No

21. Gestation Period The lengths of human pregnancies are approximately normally distributed, with mean $\mu = 266$ days and standard deviation $\sigma = 16$ days.

(a) What proportion of pregnancies lasts more than 270 days? $0.4013$
(b) What proportion of pregnancies lasts less than 250 days? $0.7590$
(c) What proportion of pregnancies lasts between 240 and 280 days? $0.1894$
(d) What is the probability that a randomly selected pregnancy lasts more than 280 days? $0.0891$
(e) What is the probability that a randomly selected pregnancy lasts no more than 245 days? $0.9671$
(f) A “very preterm” baby is one whose gestation period is less than 224 days. Are very preterm babies unusual? Yes

22. Light Bulbs General Electric manufactures a decorative Crystal Clear 60-watt light bulb that it advertises will last 1,500 hours. Suppose that the lifetimes of the light bulbs are approximately normally distributed, with a mean of 1,550 hours and a standard deviation of 57 hours.

(a) What is the probability that a randomly selected 60-watt light bulb will last at least 2,000 hours? (b) $0.1587$
23. **Manufacturing** Steel rods are manufactured with a mean length of 25 centimeter (cm). Because of variability in the manufacturing process, the lengths of the rods are approximately normally distributed, with a standard deviation of 0.07 cm.

(a) What proportion of rods has a length less than 24.9 cm?  
(b) Any rods that are shorter than 24.85 cm or longer than 25.15 cm are discarded. What proportion of rods will be discarded? 0.0324
(c) Using the results of part (b), if 5,000 rods are manufactured in a day, how many should the plant manager expect to discard? 162
(d) If an order comes in for 10,000 steel rods, how many rods should the plant manager manufacture if the order states that all rods must be between 24.9 cm and 25.1 cm? 11,804

24. **Manufacturing** Ball bearings are manufactured with a mean diameter of 5 millimeters (mm). Because of variability in the manufacturing process, the diameters of the ball bearings are approximately normally distributed, with a standard deviation of 0.02 mm.

(a) What proportion of ball bearings has a diameter more than 5.03 mm? 0.0668
(b) Any ball bearings that have a diameter less than 4.95 mm or greater than 5.05 mm are discarded. What proportion of ball bearings will be discarded? 0.0124
(c) Using the results of part (b), if 30,000 ball bearings are manufactured in a day, how many should the plant manager expect to discard? 372
(d) If an order comes in for 50,000 ball bearings, how many bearings should the plant manager manufacture if the order states that all ball bearings must be between 4.97 mm and 5.03 mm? 57,711

25. **Egg Incubation Times** The mean incubation time of fertilized chicken eggs kept at 100.5°F in a still-air incubator is 21 days. Suppose that the incubation times are approximately normally distributed with a standard deviation of 1 day.

(a) Determine the 17th percentile for incubation times of fertilized chicken eggs. 20 days
(b) Determine the incubation times that make up the middle 95% of fertilized chicken eggs. 19–23 days

26. **Reading Rates** The reading speed of sixth-grade students is approximately normal, with a mean speed of 125 words per minute and a standard deviation of 24 words per minute.

(a) What is the reading speed of a sixth-grader whose reading speed is at the 90th percentile? 156 words per minute
(b) A school psychologist wants to determine reading rates for unusual students (both slow and fast). Determine the reading rates of the middle 95% of all sixth-grade students. What are the cutoff points for unusual readers?

27. **Chips Ahoy! Cookies** The number of chocolate chips in an 18-ounce bag of Chips Ahoy! chocolate chip cookies is approximately normally distributed, with a mean of 1,262 chips and a standard deviation of 81 chips, according to a study by cadets of the U.S. Air Force Academy.

(a) Determine the 30th percentile for the number of chocolate chips in an 18-ounce bag of Chips Ahoy! cookies.
(b) Determine the number of chocolate chips in a bag of Chips Ahoy! that make up the middle 99% of bags.

28. **Wendy’s Drive-Through** Fast-food restaurants spend quite a bit of time studying the amount of time cars spend in their drive-through. Certainly, the faster the cars get through the drive-through, the more the opportunity for making money. In 2007, QSR Magazine studied drive-through times for fast-food restaurants, and Wendy’s had the best time, with a mean time a car spent in the drive-through equal to 138.5 seconds. Assume that drive-through times are normally distributed, with a standard deviation of 29 seconds. Suppose that Wendy’s wants to institute a policy at its restaurants that it will not charge any patron that must wait more than a certain amount of time for an order. Management does not want to give away free meals to more than 1% of the patrons. What time would you recommend Wendy’s advertise as the maximum wait time before a free meal is awarded? 206 seconds

29. **Speedy Lube** The time required for Speedy Lube to complete an oil change service on an automobile approximately follows a normal distribution, with a mean of 17 minutes and a standard deviation of 2.5 minutes.

(a) Speedy Lube guarantees customers that the service will take no longer than 20 minutes. If it does take longer, the customer will receive the service for half-price. What percent of customers receive the service for half-price?
(b) If Speedy Lube does not want to give the discount to more than 3% of its customers, how long should it make the guaranteed time limit? 22 minutes

30. **Putting It Together: Birth Weights** The following data represent the distribution of birth weights (in grams) for babies in which the pregnancy went full term (37 to 41 weeks). 

<table>
<thead>
<tr>
<th>Birth Weight (g)</th>
<th>Number of Live Births</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–499</td>
<td>22</td>
</tr>
<tr>
<td>500–999</td>
<td>201</td>
</tr>
<tr>
<td>1,000–1,499</td>
<td>1,645</td>
</tr>
<tr>
<td>1,500–1,999</td>
<td>9,365</td>
</tr>
<tr>
<td>2,000–2,499</td>
<td>92,191</td>
</tr>
<tr>
<td>2,500–2,999</td>
<td>569,319</td>
</tr>
<tr>
<td>3,000–3,499</td>
<td>1,387,335</td>
</tr>
<tr>
<td>3,500–3,999</td>
<td>988,011</td>
</tr>
<tr>
<td>4,000–4,499</td>
<td>255,700</td>
</tr>
<tr>
<td>4,500–4,999</td>
<td>36,766</td>
</tr>
<tr>
<td>5,000–5,499</td>
<td>3,994</td>
</tr>
</tbody>
</table>

More than 1 million skin cancers are expected to be diagnosed in the United States this year, accounting for almost half of all cancers diagnosed. The prevalence of skin cancer is attributable in part to a history of unprotected or underprotected sun exposure. Sunscreens have been shown to prevent certain types of lesions associated with skin cancer. They also protect skin against exposure to light that contributes to premature aging. As a result, sunscreen is now in moisturizers, lip balms, shampoos, hair-styling products, insect repellents, and makeup.

Consumer Reports tested 23 sunscreens and two moisturizers, all with a claimed sun-protection factor (SPF) of 15 or higher. SPF is defined as the degree to which a sunscreen protects the skin from UVB, the ultraviolet rays responsible for sunburn. (Some studies have shown that UVB, along with UVA, can increase the risk of skin cancers.) A person with untreated skin who can stay in the sun for 5 minutes before becoming sunburned should be able to stay in the sun for 15 × 5 = 75 minutes using a sunscreen rated at SPF 15.

To test whether products meet their SPF claims for UVB, we used a solar simulator (basically a sun lamp) to expose people to measured amounts of sunlight. First we determined the exposure time (in minutes) that caused each person’s untreated skin to turn pink within 24 hours. Then we applied sunscreen to new areas of skin and made the same determination. To avoid potential sources of bias, samples of the sunscreens were applied to randomly assigned sites on the subjects’ skin.

To determine the SPF rating of a sunscreen for a particular individual, the exposure time with sunscreen was divided by the exposure time without sunscreen. The following table contains the mean and standard deviation of the SPF measurements for two particular sunscreens.

<table>
<thead>
<tr>
<th>Product</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>15.5</td>
<td>1.5</td>
</tr>
<tr>
<td>B</td>
<td>14.7</td>
<td>1.2</td>
</tr>
</tbody>
</table>

(a) In designing this experiment, why is it important to obtain the exposure time without sunscreen first and then determine the exposure time with sunscreen for each person?
(b) Why is the random assignment of people and application sites to each treatment (the sunscreen) important?
(c) Assuming that the SPF ratings are approximately normal, calculate the proportion of SPF measurements that you expect to be less than 15, the advertised level of protection for each of the products.
(d) Calculate the proportion of SPF measurements that you expect to be greater than 17.5 for product A. Repeat this for product B.
(e) Calculate the proportion of SPF measurements that you expect to fall between 14.5 and 15.5 for product A. Repeat this for product B.
(f) Which product appears to be superior, A or B? Support your conclusion.

**Note to Readers:** In many cases, our test protocol and analytical methods are more complicated than described in these examples. The data and discussions have been modified to make the material more appropriate for the audience.


---

**TECHNOLOGY STEP-BY-STEP**

**TI-83/84 Plus**

**Finding Areas under the Normal Curve**

1. From the HOME screen, press 2nd VARS to access the DISTRibution menu.
2. Select 2:normalcdf(.
3. With normalcdf( on the HOME screen, type lowerbound, upperbound, μ, σ). For example, to find the area to the left of x = 35 under the normal curve with μ = 40 and σ = 10, type
   
   \[ \text{normalcdf}(-1E99, 35, 40, 10) \]
   
   and hit ENTER.
4. **Note:** When there is no lowerbound, enter \(-1E99\). When there is no upperbound, enter \(1E99\). The \(E\) shown is scientific notation; it is selected by pressing 2nd then \(\times\).

**Finding Normal Values Corresponding to an Area**

1. From the HOME screen, press 2nd VARS to access the DISTRibution menu.
2. Select 3:invNorm(.
3. With invNorm( on the HOME screen, type “area left”, μ, σ). For example, to find the normal value such that the area under the normal curve to the left of the value is 0.68, with μ = 40 and σ = 10, type
   
   \[ \text{invNorm}(0.68, 40, 10) \]
   
   and hit ENTER.
### 7.4 Assessing Normality

#### Preparing for This Section
Before getting started, review the following:
- Shape of a distribution (Section 2.2, pp. 92–93)

#### Objective
Use normal probability plots to assess normality

Suppose that we obtain a simple random sample from a population whose distribution is unknown. Many of the statistical tests that we perform on small data sets (sample size less than 30) require that the population from which the sample is drawn be normally distributed.

Up to this point, we have said that a random variable \( X \) is normally distributed, or at least approximately normal, provided the histogram of the data is symmetric and bell-shaped. This method works well for large data sets, but the shape of a histogram drawn from a small sample of observations does not always accurately represent the shape of the population. For this reason, we need additional methods for assessing the normality of a random variable \( X \) when we are looking at a small set of sample data.

### Use Normal Probability Plots to Assess Normality

A normal probability plot is a graph that plots observed data versus normal scores. A normal score is the expected z-score of the data value, assuming that the distribution of the random variable is normal. The expected z-score of an observed value depends on the number of observations in the data set.

Drawing a normal probability plot requires the following steps:

#### Drawing a Normal Probability Plot

**Step 1:** Arrange the data in ascending order.

**Step 2:** Compute \( f_i = \frac{i - 0.375}{n + 0.25} \), where \( i \) is the index (the position of the data value in the ordered list) and \( n \) is the number of observations. The expected proportion of observations less than or equal to the \( i \)th data value is \( f_i \).

**Step 3:** Find the z-score corresponding to \( f_i \) from Table V.

**Step 4:** Plot the observed values on the horizontal axis and the corresponding expected z-scores on the vertical axis.

*The derivation of this formula is beyond the scope of this text.*
Figure 43

The idea behind finding the expected z-score is that, if the data come from a population that is normally distributed, we should be able to predict the area to the left of each data value. The value of \( f_i \) represents the expected area to the left of the \( i \)th observation when the data come from a population that is normally distributed. For example, \( f_1 \) is the expected area to the left of the smallest data value, \( f_2 \) is the expected area to the left of the second-smallest data value, and so on. Figure 43 illustrates the idea.

Once we determine each \( f_i \), we find the z-scores corresponding to \( f_1 \), \( f_2 \), and so on. The smallest observation in the data set will be the smallest expected z-score, and the largest observation in the data set will be the largest expected z-score. Also, because of the symmetry of the normal curve, the expected z-scores are always paired as positive and negative values.

Values of normal random variables and their z-scores are linearly related \((x = \mu + z\sigma)\), so a plot of observations of normal variables against their expected z-scores will be linear. We conclude the following:

If sample data are taken from a population that is normally distributed, a normal probability plot of the observed values versus the expected z-scores will be approximately linear.

Normal probability plots are typically drawn using graphing calculators or statistical software. However, it is worthwhile to go through an example that demonstrates the procedure so that we can better understand the results supplied by technology.

### EXAMPLE 1

**Constructing a Normal Probability Plot**

**Problem:** The data in Table 4 represent the finishing time (in seconds) for six randomly selected races of a greyhound named Barbies Bomber in the \( \frac{5}{16} \)-mile race at Greyhound Park in Dubuque, Iowa. Is there evidence to support the belief that the variable “finishing time” is normally distributed?

**Approach:** We follow Steps 1 through 4.

**Solution**

**Step 1:** The first column in Table 5 represents the index \( i \). The second column represents the observed values in the data set, written in ascending order.

**Step 2:** The third column in Table 5 represents \( f_i = \frac{i - 0.375}{n + 0.25} \) for each observation. This value is the expected area under the normal curve to the left of the \( i \)th observation, assuming normality. For example, \( i = 1 \) corresponds to the finishing time of 31.26, and

\[
f_1 = \frac{1 - 0.375}{6 + 0.25} = 0.1
\]

So, the area under the normal curve to the left of 31.26 is 0.1, if the sample data come from a population that is normally distributed.

**Step 3:** We use Table V to find the z-scores that correspond to each \( f_i \). The expected z-scores are listed in the fourth column of Table 5. Look in Table V for the area closest to \( f_1 = 0.1 \). The expected z-score is \(-1.28\). Notice that for each negative expected z-score there is a corresponding positive expected z-score, as a result of the symmetry of the normal curve.
Section 7.4 Assessing Normality

Step 4: We plot the actual observations on the horizontal axis and the expected $z$-scores on the vertical axis. See Figure 44.

![Figure 44](image-url)

Although the normal probability plot in Figure 44 does show some curvature, it is roughly linear.* We conclude that the finishing times of Barbies Bomber in the $\frac{5}{16}$-mile race are approximately normally distributed.

Typically, normal probability plots are drawn using either a graphing calculator with advanced statistical features or statistical software. Certain software, such as MINITAB, will provide bounds that the data must lie within to support the belief that the sample data come from a population that is normally distributed.

**EXAMPLE 2**

**Assessing Normality Using Technology**

**Problem:** Using MINITAB or some other statistical software, draw a normal probability plot of the data in Table 4 and determine whether the sample data come from a population that is normally distributed.

**Approach:** We will construct a normal probability plot using MINITAB, which provides curved bounds that can be used to assess normality. If the normal probability plot is roughly linear and all the data lie within the bounds provided by the software, we have reason to believe the data come from a population that is approximately normal. The steps for constructing normal probability plots using MINITAB, Excel, or the TI-83/84 Plus graphing calculators can be found on page 411.

*In fact, the correlation between the observed value and expected $z$-score is 0.970.*
Chapter 7  The Normal Probability Distribution

The normal probability plot is roughly linear, and all the data lie within the bounds provided by MINITAB. We conclude that the sample data could come from a population that is normally distributed.

Throughout the text, we will provide normal probability plots drawn with MINITAB so that assessing normality is straightforward.

**EXAMPLE 3**

**Assessing Normality**

**Problem:** The data in Table 6 represent the time spent waiting in line (in minutes) for the Demon Roller Coaster for 100 randomly selected riders. Is the random variable “wait time” normally distributed?

**Solution:** Figure 45 shows the normal probability plot. Notice that MINITAB gives the area to the left of the expected $z$-score, rather than the $z$-score. For example, the area to the left of the expected $z$-score of $-1.28$ is 0.10. MINITAB writes 0.10 as 10 percent.

The normal probability plot is roughly linear, and all the data lie within the bounds provided by MINITAB. We conclude that the sample data could come from a population that is normally distributed.

Throughout the text, we will provide normal probability plots drawn with MINITAB so that assessing normality is straightforward.

**Table 6**

<table>
<thead>
<tr>
<th>Time (in minutes)</th>
<th>7</th>
<th>3</th>
<th>5</th>
<th>107</th>
<th>8</th>
<th>37</th>
<th>16</th>
<th>41</th>
<th>7</th>
<th>25</th>
<th>22</th>
<th>19</th>
<th>1</th>
<th>40</th>
<th>1</th>
<th>29</th>
<th>93</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>33</td>
<td>76</td>
<td>14</td>
<td>8</td>
<td>9</td>
<td>45</td>
<td>15</td>
<td>81</td>
<td>94</td>
<td>10</td>
<td>115</td>
<td>18</td>
<td>0</td>
<td>18</td>
<td>11</td>
<td>60</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>6</td>
<td>21</td>
<td>0</td>
<td>86</td>
<td>6</td>
<td>11</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>9</td>
<td>79</td>
<td>41</td>
<td>2</td>
<td>9</td>
<td>6</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>7</td>
<td>18</td>
<td>0</td>
<td>93</td>
<td>68</td>
<td>94</td>
<td>16</td>
<td>13</td>
<td>24</td>
<td>6</td>
<td>12</td>
<td>121</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>39</td>
<td>9</td>
<td>15</td>
<td>53</td>
<td>9</td>
<td>47</td>
<td>5</td>
<td>55</td>
<td>64</td>
<td>51</td>
<td>80</td>
<td>26</td>
<td>24</td>
<td>12</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>94</td>
<td>18</td>
<td>4</td>
<td>61</td>
<td>38</td>
<td>38</td>
<td>21</td>
<td>61</td>
<td>9</td>
<td>80</td>
<td>18</td>
<td>21</td>
<td>8</td>
<td>14</td>
<td>47</td>
<td>56</td>
<td></td>
</tr>
</tbody>
</table>
7.4 Assess Your Understanding

Concepts and Vocabulary

1. Explain why normal probability plots should be linear if the data are normally distributed.

2. What does $f_l$ represent?

Skill Building

In Problems 3–8, determine whether the normal probability plot indicates that the sample data could have come from a population that is normally distributed.

3. Not normal

4. Normal

5. Not normal
Chapter 7  The Normal Probability Distribution

6. Not normal

7. Normal

8. Normal

Applying the Concepts

9. Chips per Bag In a 1998 advertising campaign, Nabisco claimed that every 18-ounce bag of Chips Ahoy! cookies contained at least 1,000 chocolate chips. Brad Warner and Jim Rutledge tried to verify the claim. The following data represent the number of chips in an 18-ounce bag of Chips Ahoy! based on their study.

<table>
<thead>
<tr>
<th>Chips per Bag</th>
<th>1,087</th>
<th>1,098</th>
<th>1,103</th>
<th>1,121</th>
<th>1,132</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source: Chance 12 (1): 10–14, 1999</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) Use the following normal probability plot to determine if the data could have come from a normal distribution. Normal

(b) Determine the mean and standard deviation of the sample data. \( \bar{x} = 1247.4, \sigma = 101.0 \)

(c) Using the sample mean and sample standard deviation obtained in part (b) as estimates for the population mean and population standard deviation, respectively, draw a graph of a normal model for the distribution of chips in a bag of Chips Ahoy!.

(d) Using the normal model from part (c), find the probability that an 18-ounce bag of Chips Ahoy! selected at random contains at least 1,000 chips. 0.9929

(e) Using the normal model from part (c), determine the proportion of 18-ounce bags of Chips Ahoy! that contain between 1,200 and 1,400 chips, inclusive. 0.6153

10. Hours of TV A random sample of college students aged 18 to 24 years was obtained, and the number of hours of television watched last week was recorded.

<table>
<thead>
<tr>
<th>Hours of TV</th>
<th>36.1</th>
<th>30.5</th>
<th>2.9</th>
<th>17.5</th>
<th>21.0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>23.5</td>
<td>25.6</td>
<td>16.0</td>
<td>28.9</td>
<td>29.6</td>
</tr>
<tr>
<td></td>
<td>7.8</td>
<td>20.4</td>
<td>33.8</td>
<td>36.8</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>9.9</td>
<td>25.8</td>
<td>19.5</td>
<td>19.1</td>
<td>18.5</td>
</tr>
<tr>
<td></td>
<td>22.9</td>
<td>9.7</td>
<td>39.2</td>
<td>19.0</td>
<td>8.6</td>
</tr>
</tbody>
</table>

(a) Use the following normal probability plot to determine if the data could have come from a normal distribution. Normal
Section 7.4 Assessing Normality

411

(b) Determine the mean and standard deviation of the sample data. \( \bar{x} = 20.90, s = 10.52 \)
(c) Using the sample mean and sample standard deviation obtained in part (b) as estimates for the population mean and population standard deviation, respectively, draw a graph of a normal model for the distribution of weekly hours of television watched.
(d) Using the normal model from part (c), find the probability that a college student aged 18 to 24 years, selected at random, watches between 20 and 35 hours of television each week. 0.4457
(e) Using the normal model from part (c), determine the proportion of college students aged 18 to 24 years who watch more than 40 hours of television per week. 0.0344

In Problems 11–14, use a normal probability plot to assess whether the sample data could have come from a population that is normally distributed.

11. O-Ring Thickness A random sample of O-rings was obtained, and the wall thickness (in inches) of each was recorded. Normal

12. Customer Service A random sample of weekly work logs at an automobile repair station was obtained, and the average number of customers per day was recorded. Normal

13. School Loans A random sample of 20 undergraduate students receiving student loans was obtained, and the amount of their loans for the 2007–2008 school year was recorded. Not normal

14. Memphis Snowfall A random sample of 25 years between 1890 and 2007 was obtained, and the amount of snowfall, in inches, for Memphis was recorded. Not normal

TECHNOLOGY STEP-BY-STEP

Normal Probability Plots

3. In the Graph variables cell, enter the column that contains the raw data. Make sure Distribution is set to Normal. Click OK.

Excel
1. Install Data Desk XL.
2. Enter the raw data into column A.
3. Select the DDXL menu. Highlight Charts and Plots.
4. In the pulldown menu, select Normal Probability Plot. Drag the column containing the data to the Quantitative Variable cell and click OK. If the first row contains the variable name, check the “First row is variable name” box.
Historical Note
The normal approximation to the binomial was discovered by Abraham de Moivre in 1733. With the advance of computing technology, its importance has been diminished.

Approximate Binomial Probabilities Using the Normal Distribution

In Section 6.2, we discussed the binomial probability distribution. Below, we remind you of the criteria for a probability experiment to be a binomial experiment.

Criteria for a Binomial Probability Experiment

A probability experiment is said to be a binomial experiment if all the following are true:

1. The experiment is performed \( n \) independent times. Each repetition of the experiment is called a trial. Independence means that the outcome of one trial will not affect the outcome of the other trials.
2. For each trial, there are two mutually exclusive outcomes—success or failure.
3. The probability of success, \( p \), is the same for each trial of the experiment.

The binomial probability formula can be used to compute probabilities of events in a binomial experiment. When there is a large number of trials of a binomial experiment, the binomial probability formula can be difficult to use. For example, for 500 trials of a binomial experiment, we wish to compute the probability of 400 or more successes. Using the binomial probability formula requires that we compute the following probabilities:

\[
P(X \geq 400) = P(400) + P(401) + \cdots + P(500)
\]

This would be time consuming to compute by hand! Fortunately, we have an alternate means for approximating binomial probabilities, provided that certain conditions are met.

Recall, for a fixed \( p \), as the number of trials \( n \) in a binomial experiment increases, the probability histogram becomes more nearly symmetric and bell-shaped (see page 353). We restate the conclusion here.

For a fixed \( p \), as the number of trials \( n \) in a binomial experiment increases, the probability distribution of the random variable \( X \) becomes more nearly symmetric and bell-shaped. As a rule of thumb, if \( np(1 - p) \geq 10 \), the probability distribution will be approximately symmetric and bell-shaped.

Because of this result, we might be inclined to think that binomial probabilities can be approximated by the area under a normal curve, provided that \( np(1 - p) \geq 10 \). This intuition is correct.
Section 7.5 The Normal Approximation to the Binomial Probability Distribution

The Normal Approximation to the Binomial Probability Distribution

If \( np(1 - p) \geq 10 \), the binomial random variable \( X \) is approximately normally distributed, with mean \( \mu_X = np \) and standard deviation \( \sigma_X = \sqrt{np(1 - p)} \).

Figure 48 shows a probability histogram for the binomial random variable \( X \), with \( n = 40 \) and \( p = 0.5 \), and a normal curve, with \( \mu_X = np = 40(0.5) = 20 \) and standard deviation \( \sigma_X = \sqrt{np(1 - p)} = \sqrt{40(0.5)(0.5)} = \sqrt{10} \). Notice that \( np(1 - p) = 40(0.5)(1 - 0.5) = 10 \).

![Binomial Histogram, n = 40, p = 0.5](image1)

We know from Section 6.2 that the area of the rectangle corresponding to \( x = 18 \) represents \( P(18) \). The width of each rectangle is 1, so the rectangle extends from \( x = 17.5 \) to \( x = 18.5 \). The area under the normal curve from \( x = 17.5 \) to \( x = 18.5 \) is approximately equal to the area of the rectangle corresponding to \( x = 18 \). Therefore, the area under the normal curve between \( x = 17.5 \) and \( x = 18.5 \) is approximately equal to \( P(18) \), where \( X \) is a binomial random variable with \( n = 40 \) and \( p = 0.5 \). We add and subtract 0.5 from \( x = 18 \) as a correction for continuity, because we are using a continuous density function to approximate a discrete probability.

Suppose that we want to approximate \( P(X \leq 18) \). Figure 49 illustrates the situation.

![Binomial Histogram, n = 40, p = 0.5](image2)

To approximate \( P(X \leq 18) \), we compute the area under the normal curve for \( x < 18.5 \). Do you see why?

If we want to approximate \( P(X \geq 18) \), we compute \( P(X \geq 17.5) \). Do you see why? Table 7 summarizes how to use the correction for continuity.
A question remains, however. What do we do if the probability is of the form \( P(X > a), P(X < a), \) or \( P(a < X < b)? \) The solution is to rewrite the inequality in a form with \( \leq \) or \( \geq \). For example, \( P(X > 4) = P(X \geq 5) \) and \( P(X < 4) = P(X \leq 3) \) for binomial random variables, because the values of the random variables must be whole numbers.

**Table 7**

<table>
<thead>
<tr>
<th>Exact Probability Using Binomial</th>
<th>Approximate Probability Using Normal</th>
<th>Graphical Depiction</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P(a) )</td>
<td>( P(a - 0.5 \leq X \leq a + 0.5) )</td>
<td><img src="image1" alt="Graph" /></td>
</tr>
<tr>
<td>( P(X \leq a) )</td>
<td>( P(X \leq a + 0.5) )</td>
<td><img src="image2" alt="Graph" /></td>
</tr>
<tr>
<td>( P(X \geq a) )</td>
<td>( P(X \geq a - 0.5) )</td>
<td><img src="image3" alt="Graph" /></td>
</tr>
<tr>
<td>( P(a \leq X \leq b) )</td>
<td>( P(a - 0.5 \leq X \leq b + 0.5) )</td>
<td><img src="image4" alt="Graph" /></td>
</tr>
</tbody>
</table>

**EXAMPLE 1**

**The Normal Approximation to a Binomial Random Variable**

**Problem:** According to the *American Red Cross*, 7% of people in the United States have blood type O-negative. What is the probability that, in a simple random sample of 500 people in the U.S., fewer than 30 have blood type O-negative?

**Approach**

**Step 1:** We verify that this is a binomial experiment.

**Step 2:** Computing the probability by hand would be very tedious. Verify \( np(1 - p) \geq 10 \). Then we will know that the condition for using the normal distribution to approximate the binomial distribution is met.

**Step 3:** Approximate \( P(X < 30) = P(X \leq 29) \) by using the normal approximation to the binomial distribution.

**Solution**

**Step 1:** There are 500 independent trials with each trial having a probability of success equal to 0.07. This is a binomial experiment.

**Step 2:** We verify \( np(1 - p) \geq 10 \).

\[
np(1 - p) = 500(0.07)(0.93) = 32.55 \geq 10
\]

We can use the normal distribution to approximate the binomial distribution.
Step 3: We wish to know the probability that fewer than 30 people in the sample have blood type O-negative; that is, we wish to know \( P(X < 30) = P(X \leq 29) \). This is approximately equal to the area under the normal curve to the left of \( x = 29.5 \), with \( \mu_X = np = 500(0.07) = 35 \) and \( \sigma_X = \sqrt{np(1 - p)} = \sqrt{500(0.07)(1 - 0.07)} = \sqrt{32.55} \approx 5.71 \). See Figure 50. We convert \( x = 29.5 \) to a z-score.

\[
z = \frac{29.5 - 35}{\sqrt{32.55}} = -0.96
\]

From Table V, we find that the area to the left of \( z = -0.96 \) is 0.1685. Therefore, the approximate probability that fewer than 30 people will have blood type O-negative is 0.1685 = 16.85%.

Using the \( \text{binomcdf} \) command on a TI-84 Plus graphing calculator, we find that the exact probability is 0.1678. See Figure 51. The approximate result is close indeed!

**EXAMPLE 2**

A Normal Approximation to the Binomial

**Problem:** According to the Federal Communications Commission, 70% of all U.S. households have cable television. Erica selects a random sample of 1,000 households in DuPage County and finds that 734 of them have cable.

(a) Assuming that 70% of households have cable, what is the probability of obtaining a random sample of at least 734 households with cable from a sample of size 1,000?

(b) Does the result from part (a) contradict the FCC information? Explain.

**Approach:** This is a binomial experiment with \( n = 1,000 \) and \( p = 0.70 \). Erica needs to determine the probability of obtaining a random sample of at least 734 households with cable from a sample of size 1,000, assuming 70% of households have cable. Computing this using the binomial probability formula would be difficult, so Erica will compute the probability using the normal approximation to the binomial, since \( np(1 - p) = 1,000(0.70)(0.30) = 210 \geq 10 \). We approximate \( P(X \geq 734) \) by computing the area under the standard normal curve to the right of \( x = 733.5 \), with \( \mu_X = np = 1,000(0.70) = 700 \) and \( \sigma_X = \sqrt{np(1 - p)} = \sqrt{1,000(0.70)(1 - 0.70)} = \sqrt{210} \approx 14.491 \).

**Solution**

(a) Figure 52 shows the area we wish to compute. We convert \( x = 733.5 \) to a z-score.

\[
z = \frac{733.5 - 700}{\sqrt{210}} = 2.31
\]

The area under the standard normal curve to the right of \( z = 2.31 \) is \( 1 - 0.9896 = 0.0104 \). There is a 1.04% probability of obtaining 734 or more households with cable from a sample of 1,000 households, assuming that the percentage of households with cable is 70%.

(b) The result from part (a) means that about 1 sample in every 100 samples will have 734 or more households with cable if the true proportion is 0.7. Erica is not inclined to believe that her sample is one of the 1 in 100. She would conclude that the proportion of households in DuPage County with cable is higher than 0.70.
Chapter 7 The Normal Probability Distribution

7.5 ASSESS YOUR UNDERSTANDING

Concepts and Vocabulary
1. List the conditions required for a binomial experiment.
2. Under what circumstances can the normal distribution be used to approximate binomial probabilities?
3. Why must we use a correction for continuity when using the normal distribution to approximate binomial probabilities?
4. True or False: Suppose that \( X \) is a binomial random variable. To approximate \( P(3 \leq X < 7) \) using the normal probability distribution, we compute \( P(3.5 \leq X < 7.5) \). False

Skill Building

In Problems 5–14, a discrete random variable is given. Assume the probability of the random variable will be approximated using the normal distribution. Describe the area under the normal curve that will be computed. For example, if we wish to compute the probability of finding at least five defective items in a shipment, we will compute \( P(X \geq 5) \).

5. The probability that at least 40 households have a gas stove
   Right of \( x = 39.5 \)
6. The probability of no more than 20 people who want to see Roe v. Wade overturned
   Left of \( x = 20.5 \)
7. The probability that exactly eight defective parts are in the shipment
   Between \( x = 7.5 \) and \( x = 8.5 \)
8. The probability that exactly 12 students pass the course
   Between \( x = 7.5 \) and \( x = 8.5 \)
9. The probability that the number of people with blood type O-negative is between 18 and 24, inclusive
10. The probability that the number of tornadoes that occur in the month of May is between 30 and 40, inclusive
11. The probability that more than 20 people want to see the marriage tax penalty abolished
   Right of \( x = 20.5 \)
12. The probability that fewer than 40 households have a pet
13. The probability that more than 500 adult Americans support a bill proposing to extend daylight savings time
14. The probability that fewer than 35 people support the privatization of Social Security
   Left of \( x = 34.5 \)

In Problems 15–20, compute \( P(x) \) using the binomial probability formula. Then determine whether the normal distribution can be used as an approximation for the binomial distribution. If so, approximate \( P(x) \) and compare the result to the exact probability.

15. \( n = 60, p = 0.4, x = 20 \)
   \( 0.0616; 0.0616 \)
16. \( n = 80, p = 0.15, x = 18 \)
   \( 0.0221; 0.0216 \)
17. \( n = 40, p = 0.25, x = 30 \)
   \(< 0.0001; \text{cannot be used}\)
18. \( n = 100, p = 0.05, x = 50 \)
   \(< 0.0001; \text{cannot be used}\)
19. \( n = 75, p = 0.75, x = 60 \)
   \( 0.0677; 0.0630 \)
20. \( n = 85, p = 0.8, x = 70 \)
   \( 0.0970; 0.0926 \)

Applying the Concepts

21. On-Time Flights According to American Airlines, Flight 215 from Orlando to Los Angeles is on time 90% of the time. Randomly select 150 flights and use the normal approximation to the binomial to
   (a) approximate the probability that exactly 130 flights are on time. 0.0444
   (b) approximate the probability that at least 130 flights are on time. 0.9932
   (c) approximate the probability that fewer than 125 flights are on time. 0.0021
   (d) approximate the probability that between 125 and 135 flights inclusive, are on time. 0.5536

22. Smokers According to Information Please Almanac, 80% of adult smokers started smoking before they were 18 years old. Suppose 100 smokers 18 years old or older, are randomly selected. Use the normal approximation to the binomial to
   (a) approximate the probability that exactly 80 of them started smoking before they were 18 years old. 0.1034
   (b) approximate the probability that at least 80 of them started smoking before they were 18 years old. 0.5917
   (c) approximate the probability that fewer than 70 of them started smoking before they were 18 years old. 0.0045
   (d) approximate the probability that between 70 and 90 of them, inclusive, started smoking before they were 18 years old. 0.9914

23. Migraine Sufferers In clinical trials of a medication whose purpose is to reduce the pain associated with migraine headaches, 2% of the patients in the study experienced weight gain as a side effect. A random sample of 600 users of this medication is obtained. Use the normal approximation to the binomial to
   (a) approximate the probability that exactly 20 will experience weight gain as a side effect. 0.0077
   (b) approximate the probability that 20 or fewer will experience weight gain as a side effect. 0.9923
   (c) approximate the probability that 22 or more patients will experience weight gain as a side effect. 0.0028
   (d) approximate the probability that between 20 and 30 patients, inclusive, will experience weight gain as a side effect. 0.0143

24. Murder by Firearms According to the Uniform Crime Report, 2005, 67.8% of murders are committed with a firearm. Suppose 50 murders are randomly selected. Use the normal approximation to the binomial to
   (a) approximate the probability that exactly 40 murders are committed using a firearm. 0.0227
   (b) approximate the probability that at least 35 murders are committed using a firearm. 0.4286
   (c) approximate the probability that fewer than 25 murders are committed using a firearm. 0.0022
   (d) approximate the probability that between 20 and 35 murders, inclusive, are committed using a firearm. 0.5926

25. Males Living at Home According to the Current Population Survey (Internet release date: September 15, 2004), 55% of males between the ages of 18 and 24 years lived at home in
   (a) approximate the probability that between 29 and 40 of them, inclusive, lived at home. 0.5826

8. Between \( x = 11.5 \) and \( x = 12.5 \)
9. Between \( x = 17.5 \) and \( x = 24.5 \)
12. Left of \( x = 39.5 \)
13. Right of \( x = 500.5 \)
2003. (Unmarried college students living in a dorm are counted as living at home.) A survey is administered at a community college to 200 randomly selected male students between the ages of 18 and 24 years, and 130 of them respond that they live at home.

(a) Approximate the probability that such a survey will result in at least 130 of the respondents living at home under the assumption that the true percentage is 55%. 0.0028

(b) Does the result from part (a) contradict the results of the Current Population Survey? Explain. Yes

26. Females Living at Home According to the Current Population Survey (Internet release date: September 15, 2004), 46% of females between the ages of 18 and 24 years lived at home in 2003. (Unmarried college students living in a dorm are counted as living at home.) A survey is administered at a community college to 200 randomly selected female students between the ages of 18 and 24 years, and 110 of them respond that they live at home.

(a) Approximate the probability that such a survey will result in at least 110 of the respondents living at home under the assumption that the true percentage is 46%. 0.0066

(b) Does the result from part (a) contradict the results of the Current Population Survey? Explain. Yes

27. Boys Are Preferred In a Gallup poll conducted June 11–14, 2007, 37% of survey respondents said that, if they only had one child, they would prefer the child to be a boy. You conduct a survey of 150 randomly selected students on your campus and find that 80 of them would prefer a boy.

(a) Approximate the probability that, in a random sample of 150 students, at least 75 would prefer a boy, assuming the true percentage is 37%. 0.0007

(b) Does this result contradict the Gallup poll? Explain. Yes

28. Liars According to a USA Today “Snapshot,” 3% of Americans surveyed lie frequently. You conduct a survey of 500 college students and find that 20 of them lie frequently.

(a) Compute the probability that, in a random sample of 500 college students, at least 20 lie frequently, assuming the true percentage is 3%. 0.1190

(b) Does this result contradict the USA Today “Snapshot”? Explain. No

## Chapter 7 Review

### Summary

In this chapter, we introduced continuous random variables and the normal probability density function. A continuous random variable is said to be approximately normally distributed if a histogram of its values is symmetric and bell-shaped. In addition, we can draw normal probability plots that are based on expected \( z \)-scores. If these normal probability plots are approximately linear, we say the distribution of the random variable is approximately normal. The area under the normal density function can be used to find proportions, probabilities, or percentiles for normal random variables. Also, we can find the value of a normal random variable that corresponds to a specific proportion, probability, or percentile.

If \( X \) is a binomial random variable with \( np(1 - p) \geq 10 \), we can use the area under the normal curve to approximate the probability of a binomial random variable. The parameters of the normal curve are \( \mu_X = np \) and \( \sigma_X = \sqrt{np(1 - p)} \), where \( n \) is the number of trials of the binomial experiment and \( p \) is the probability of success.

### Vocabulary

- Uniform probability distribution (p. 373)
- Probability density function (p. 373)
- Normal curve (p. 375)
- Normally distributed (p. 375)
- Normal probability distribution (p. 375)
- Inflection points (p. 375)
- Normal density curve (p. 376)
- Model (p. 377)
- Normal probability density function (p. 378)
- Standard normal distribution (p. 379)
- Standard normal curve (p. 385)
- Normal probability plot (p. 405)
- Normal score (p. 405)
- Trial (p. 412)
- Normal approximation to the binomial distribution (p. 413)
- Correction for continuity (p. 413)

### Formulas

**Standardizing a Normal Random Variable**

\[
 z = \frac{x - \mu}{\sigma}
\]

**Finding the Score**

\[
x = \mu + z\sigma
\]


## Review Exercises

1. Use the figure to answer the questions that follow:

   ![Normal Distribution Curve](image)

   (a) What is μ? 60
   (b) What is σ? 10
   (c) Suppose that the area under the normal curve to the right of x = 75 is 0.0668. Provide two interpretations for this area.
   (d) Suppose that the area under the normal curve between x = 50 and x = 75 is 0.7745. Provide two interpretations for this area.

2. A random variable X is approximately normally distributed, with μ = 20 and σ = 4.
   (a) Compute \( z_1 = \frac{x_1 - \mu}{\sigma} \) for \( x_1 = 18 \). \(-0.5\)
   (b) Compute \( z_2 = \frac{x_2 - \mu}{\sigma} \) for \( x_2 = 21 \). \(0.25\)
   (c) The area under the normal curve between \( x_1 = 18 \) and \( x_2 = 21 \) is 0.2912. What is the area between \( z_1 \) and \( z_2 \)? \(0.2912\)

In Problems 3 and 4, draw a standard normal curve and shade the area indicated. Then find the area of the shaded region.

3. The area to the left of \( z = -1.04 \). \(0.1492\)
4. The area between \( z = -0.34 \) and \( z = 1.03 \). \(0.4816\)

In Problems 5–7, find the indicated probability of the standard normal random variable \( Z \).

5. \( P(Z < 1.19) \) \(0.8830\)
6. \( P(Z \geq 1.61) \) \(0.0537\)
7. \( P(-1.21 < Z \leq 2.28) \) \(0.8756\)
8. Find the \( z \)-score such that the area to the right of the \( z \)-score is 0.483. \(0.04\)
9. Find the \( z \)-scores that separate the middle 92% of the data from the area in the tails of the standard normal distribution. \(-1.75, 1.75\)
10. Find the value of \( z_{0.20} \). \(0.84\)

In Problems 11–13, draw the normal curve with the parameters indicated. Then find the probability of the random variable \( X \). Shade the area that represents the probability.

11. \( \mu = 50, \sigma = 6, P(X > 55) \) \(0.2033\)
12. \( \mu = 30, \sigma = 5, P(X \leq 23) \) \(0.0800\)
13. \( \mu = 70, \sigma = 10, P(65 < X < 85) \) \(0.6247\)

14. **Tire Wear** Suppose that Dunlop Tire manufactures tires having the property that the mileage the tire lasts approximately follows a normal distribution with mean 70,000 miles and standard deviation 4,400 miles.
   (a) What proportion of the tires will last at least 75,000 miles? \(0.1271\)
   (b) Suppose that Dunlop warrants the tires for 60,000 miles. What proportion of the tires will last 60,000 miles or less? \(0.016\)
   (c) What is the probability that a randomly selected Dunlop tire lasts between 65,000 and 80,000 miles? \(0.8613\)
   (d) Suppose that Dunlop wants to warrant no more than 2% of its tires. What mileage should the company advertise as its warranty mileage? \(60,980\) miles

15. **Wechsler Intelligence Scale** The Wechsler Intelligence Scale for Children is approximately normally distributed, with mean 100 and standard deviation 15.
   (a) What is the probability that a randomly selected test taker will score above 125? \(0.0475\)
   (b) What is the probability that a randomly selected test taker will score below 90? \(0.2514\)
   (c) What proportion of test takers will score between 110 and 140? \(0.2476\)
   (d) If a child is randomly selected, what is the probability that she scores above 150? \(0.0004\)
   (e) What intelligence score will place a child in the 98th percentile? \(131\)
(f) If normal intelligence is defined as scoring in the middle 95% of all test takers, figure out the scores that differentiate normal intelligence from abnormal intelligence. 71–129

16. Major League Baseballs  According to major league baseball rules, the ball must weigh between 5 and 5.25 ounces. A factory produces baseballs whose weights are approximately normally distributed, with mean 5.11 ounces and standard deviation 0.062 ounces.

Source: www.baseball-almanac.com

(a) What proportion of the baseballs produced by this factory are too heavy for use by major league baseball? 0.0119
(b) What proportion of the baseballs produced by this factory are too light for use by major league baseball? 0.0384
(c) What proportion of the baseballs produced by this factory can be used by major league baseball? 0.9497
(d) If 8,000 baseballs are ordered, how many baseballs should be manufactured, knowing that some will need to be discarded? 8,424

17. America Reads  According to a Gallup poll conducted May 20–22, 2005, 46% of Americans 18 years old or older stated that they had read at least six books (fiction and non-fiction) within the past year. You conduct a random sample of 250 Americans 18 years old or older.

(a) Verify that the conditions for using the normal distribution to approximate the binomial distribution are met.
(b) Approximate the probability that exactly 125 read at least six books within the past year. Interpret this result. 0.0214
(c) Approximate the probability that fewer than 120 read at least six books within the past year. Interpret this result. 0.7157
(d) Approximate the probability that at least 140 read at least six books within the past year. Interpret this result. 0.0009
(e) Approximate the probability that between 100 and 120, inclusive, read at least six books within the past year. Interpret this result. 0.7336

In Problems 18 and 19, a normal probability plot of a simple random sample of data from a population whose distribution is unknown was obtained. Given the normal probability plot, is there reason to believe the population is normally distributed?

18. Normal

19. Not normal

20. Density of Earth  In 1798, Henry Cavendish obtained 27 measurements of the density of Earth, using a torsion balance. The following data represent his estimates, given as a multiple of the density of water. Is it reasonable to conclude that the sample data come from a population that is normally distributed? No

<table>
<thead>
<tr>
<th>5.50</th>
<th>5.57</th>
<th>5.42</th>
<th>5.61</th>
<th>5.53</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.47</td>
<td>4.88</td>
<td>5.62</td>
<td>5.63</td>
<td>4.07</td>
</tr>
<tr>
<td>5.29</td>
<td>5.34</td>
<td>5.26</td>
<td>5.44</td>
<td>5.55</td>
</tr>
<tr>
<td>5.34</td>
<td>5.30</td>
<td>5.36</td>
<td>5.79</td>
<td>5.29</td>
</tr>
<tr>
<td>5.10</td>
<td>5.86</td>
<td>5.58</td>
<td>5.27</td>
<td>5.85</td>
</tr>
<tr>
<td>5.65</td>
<td>5.39</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


21. Creative Thinking  According to a USA Today “Snapshot,” 20% of adults surveyed do their most creative thinking while driving. You conduct a survey of 250 adults and find that 30 do their most creative thinking while driving.

(a) Compute the probability that, in a random sample of 250 adults, 30 or fewer do their most creative thinking while driving. 0.0010
(b) Does this result contradict the USA Today “Snapshot”? Explain. Yes

22. A continuous random variable $X$ is uniformly distributed with $0 \leq X \leq 20$.

(a) Draw a graph of the uniform density function.
(b) What is $P(0 \leq X \leq 5)$? 0.25
(c) What is $P(10 \leq X \leq 18)$? 0.4

23. List the properties of the standard normal curve.

24. Explain how to use a normal probability plot to assess normality.
Chapter 7 The Normal Probability Distribution

CHAPTER TEST

1. Use the figure to answer the questions that follow:

(a) What is $\mu$? 7
(b) What is $\sigma$? 2
(c) Suppose that the area under the normal curve to the left of $x = 10$ is 0.9332. Provide two interpretations for this area.
(d) Suppose that the area under the normal curve between $x = 5$ and $x = 8$ is 0.5328. Provide two interpretations for this area.

2. A random variable $X$ is approximately normally distributed with $\mu = 50$ and $\sigma = 8$.
(a) Compute $z_1 = \frac{x_1 - \mu}{\sigma}$ for $x_1 = 48$. $-0.25$
(b) Compute $z_2 = \frac{x_2 - \mu}{\sigma}$ for $x_2 = 60$. 1.25
(c) The area under the normal curve between $x_1 = 48$ and $x_2 = 60$ is 0.4931. What is the area between $z_1$ and $z_2$? 0.4931

3. Draw a standard normal curve and shade the area to the right of $z = 2.04$. Then find the area of the shaded region. 0.0207

4. Find $P(0.21 < Z < 1.69)$. 0.3713

5. Find the $z$-scores that separate the middle 88% of the data from the area in the tails of the standard normal distribution. $-1.555, 1.555$

6. Find the value of $z_{0.04}$. 1.75

7. (a) Draw a normal curve with $\mu = 20$ and $\sigma = 3$.
(b) Shade the region that represents $P(22 \leq X \leq 27)$ and find the probability. 0.2415

8. Suppose that the talk time on the Apple iPhone is approximately normally distributed with mean 7 hours and standard deviation 0.8 hour.
(a) What proportion of the time will a fully charged iPhone last at least 6 hours? 0.8944
(b) What is the probability a fully charged iPhone will last less than 5 hours? 0.0062
(c) What talk time would represent the cutoff for the top 5% of all talk times? 8.5 hours
(d) Would it be unusual for the phone to last more than 9 hours? Why? Yes

9. The waist circumference of males 20 to 29 years old is approximately normally distributed, with mean 92.5 cm and standard deviation 13.7 cm. 
(a) Use the normal model to determine the proportion of 20- to 29-year-old males whose waist circumference is less than 100 cm. 0.7088
(b) What is the probability that a randomly selected 20- to 29-year-old male has a waist circumference between 80 and 100 cm? 0.5274
(c) Determine the waist circumferences that represent the middle 90% of all waist circumferences. 70 to 115 cm
(d) Determine the waist circumference that is at the 10th percentile. 75 cm

10. In a poll conducted by the Gallup organization August 13–16, 2007, 16% of adult, employed Americans were dissatisfied with the amount of their vacation time. You conduct a survey of 500 adult, employed Americans.
(a) Approximate the probability that exactly 100 are dissatisfied with their amount of vacation time. 0.0025
(b) Approximate the probability that less than 60 are dissatisfied with the amount of their vacation time. 0.0062

11. Jane obtained a random sample of 15 college students and asked how many hours they studied last week. Is it reasonable to believe that hours studied is normally distributed based on the following normal probability plot? No; not normal

12. A continuous random variable $X$ is uniformly distributed with $10 \leq X \leq 50$.
(a) Draw a graph of the uniform density function.
(b) What is $P(20 \leq X \leq 30)$? 0.25
(c) What is $P(X < 15)$? 0.125
Join the Club

You are interested in starting your own MENSA-type club. To qualify for the club, the potential member must have intelligence that is in the top 20% of all people. The problem that you face is that you do not have a baseline for measuring what qualifies as a top 20% score. To gather data, you must obtain a random sample of at least 25 volunteers to take an online intelligence test. Many online intelligence tests are available, but you need to make sure that the test will supply scored exams. One suggested site is www.queendom.com/tests/iq/classical_iq_r2_access.html.

Once you have obtained your sample of at least 25 test scores, answer the following questions.
(a) What is the mean test score? What is the standard deviation of the test scores?
(b) Do the sample data come from a population that is normally distributed? How do you know this?
(c) Assuming that the sample data come from a population that is normally distributed, determine the test score that would be required to join your club. That is, determine the test score that serves as a cutoff point for the top 20%. You can use this score to determine which potential members may join!

A Tale of Blood Chemistry and Health

Abby Tudor recently turned 40 years old. Her knees ache, and she often feels short of breath during exercise. She is experiencing fatigue and often feels that she is going in slow motion. Periodic dizziness plagues her during most days. According to the drugstore machine, her blood pressure is elevated. Her family has a history of cardiac disease, with both of her parents having experienced heart attacks. Additionally, an aunt on her mother’s side has diabetes. Hypothyroidism also runs throughout her immediate family. Abby is approximately 20 pounds overweight. She has tried various diet and exercise programs in an attempt to lose weight. Her results have been disappointing.

With the advice of her physician, she scheduled an appointment for a full physical exam, including a complete blood workup. Her doctor is particularly interested in the level of various blood components that might shed some light on Abby’s reported symptoms. Specifically, he wants to examine her white blood cell count, red blood cell count, hemoglobin, and hematocrit figures for indications of infection or anemia. Her serum glucose level will provide information concerning the possibility of the onset of diabetes. Cholesterol and triglyceride levels will provide insight into potential cardiac problems. Additionally, the possibility of hypothyroidism will be investigated by examining Abby’s serum TSH level.

As instructed, two weeks before her doctor’s appointment, she reported to her doctor’s lab for a blood test. She confirmed for the lab technician that she had fasted for the 12 hours immediately preceding the exam.
During her physical, Abby’s doctor went over the blood test report with her. He expressed concern over some of the results, but Abby was not convinced that she had a problem. Additional blood tests were not a viable option because it takes a good deal of time to get a sample of readings and they are expensive. At the time of her doctor’s appointment, she was unwilling to accept the offered prescriptions. She chose to do a little research before committing herself to any drug regimen.

Her research revealed that many medical measurements, such as cholesterol, are normally distributed in healthy populations. Unfortunately, the lab report did not provide the means and standard deviations necessary for Abby to calculate the various probabilities of interest. However, the report did provide the appropriate reference intervals. Assuming that the reference intervals represent the range of values for each blood component for a healthy adult population, it is possible to estimate the various means and standard deviations for this population. Abby estimated each mean by taking the midpoint of its reference interval. Using the Range Rule of Thumb ($\sigma \approx \text{range}/4$), standard deviations were estimated by dividing the reference interval range by 4. The following table lists Abby’s blood test results, as well as the mean and standard deviation for a number of blood test components for the population of normal healthy adults.

For any blood component measurement that was below its population mean, Abby decided to calculate the probability that she would get a test value less than or equal to the value obtained, given that she was a member of the healthy population. For example, her HDL cholesterol reading (42 mg/dL) was below the mean of the healthy population (92.5 mg/dL), so she calculated the following probability:

$$P(X \leq 42 \text{ mg/dL})$$

<table>
<thead>
<tr>
<th>Abby’s Blood Test Component</th>
<th>Unit</th>
<th>Standard Mean</th>
<th>Abby’s Absolute Deviation</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>White blood cell count</td>
<td>$10^9$/mL</td>
<td>7.25</td>
<td>1.625</td>
<td>5.3</td>
</tr>
<tr>
<td>Red blood cell count</td>
<td>$10^9$/mL</td>
<td>4.85</td>
<td>0.375</td>
<td>4.62</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>g/dL</td>
<td>14.75</td>
<td>1.125</td>
<td>14.6</td>
</tr>
<tr>
<td>Hematocrit</td>
<td>%</td>
<td>43.0</td>
<td>3.5</td>
<td>41.7</td>
</tr>
<tr>
<td>Glucose, serum</td>
<td>mg/dL</td>
<td>87.0</td>
<td>11.0</td>
<td>95.0</td>
</tr>
<tr>
<td>Creatine, serum</td>
<td>mg/dL</td>
<td>1.00</td>
<td>0.25</td>
<td>0.8</td>
</tr>
<tr>
<td>Sodium, serum</td>
<td>mEq/L</td>
<td>141.5</td>
<td>3.25</td>
<td>143.0</td>
</tr>
<tr>
<td>Potassium, serum</td>
<td>mEq/L</td>
<td>4.5</td>
<td>0.5</td>
<td>5.1</td>
</tr>
<tr>
<td>Chloride, serum</td>
<td>mEq/L</td>
<td>102.5</td>
<td>3.25</td>
<td>100.0</td>
</tr>
<tr>
<td>Carbon dioxide, total</td>
<td>mEq/L</td>
<td>26.0</td>
<td>3.0</td>
<td>25.0</td>
</tr>
<tr>
<td>Calcium, serum</td>
<td>mg/dL</td>
<td>9.55</td>
<td>0.525</td>
<td>10.1</td>
</tr>
<tr>
<td>Total cholesterol</td>
<td>mg/dL</td>
<td>149.5</td>
<td>24.75</td>
<td>253.0</td>
</tr>
<tr>
<td>Triglycerides</td>
<td>mg/dL</td>
<td>99.5</td>
<td>49.75</td>
<td>150.0</td>
</tr>
<tr>
<td>HDL cholesterol</td>
<td>mg/dL</td>
<td>92.5</td>
<td>28.75</td>
<td>42.0</td>
</tr>
<tr>
<td>LDL cholesterol</td>
<td>mg/dL</td>
<td>64.5</td>
<td>32.25</td>
<td>181.0</td>
</tr>
<tr>
<td>LDL/HDL ratio</td>
<td>Ratio</td>
<td>1.8</td>
<td>0.72</td>
<td>4.3</td>
</tr>
<tr>
<td>TSH, high sensitivity, serum</td>
<td>mcIU/mL</td>
<td>2.925</td>
<td>1.2875</td>
<td>3.15</td>
</tr>
</tbody>
</table>

*Population means and standard deviations were estimated from the reference intervals derived from an actual blood test report provided by TA LabCorp, Tampa, Florida. Means were estimated by taking the midpoints of the reference intervals. Standard deviations were estimated by dividing the reference interval ranges by 4. Test results attributed to Abby Tudor are actual results obtained from an anonymous patient.
Similarly, for any blood component measurement reading exceeding its population mean, Abby decided to calculate the probability that she would get a test value greater than or equal to the value obtained, given that she was a member of the healthy population. For example, her LDL cholesterol value (181 mg/dL) exceeds the mean of the healthy population (64.5 mg/dL), so she calculated the following probability:

\[ P(X \geq 181 \text{ mg/dL}) \]

To help her interpret the calculated probabilities, Abby decided to be concerned about only those blood components that had a probability of less than 0.025. By choosing this figure, she is acknowledging that it is unlikely that she could have such an extreme blood component reading and still be part of the healthy population.

1. The reference interval for HDL cholesterol is 35 to 150 mg/dL. Use this information to confirm the mean and standard deviation provided for this blood component.

2. Using Abby’s criteria and the means and standard deviations provided in her blood test report, determine which blood components should be a cause of concern for Abby. Write up a summary report of your findings. Be sure to include a discussion concerning your assumptions and any limitations to your conclusions.