## Chapter Outline

1.1 Approximating Areas
1.2 The Definite Integral
1.3 The Fundamental Theorem of Calculus
1.4 Integration Formulas and the Net Change Theorem
1.5 Substitution
1.6 Integrals Involving Exponential and Logarithmic Functions
1.7 Integrals Resulting in Inverse Trigonometric Functions


**Figure 1.1** Iceboating is a popular winter sport in parts of the northern United States and Europe. (credit: modification of work by Carter Brown, Flickr)

Iceboats are a common sight on the lakes of Wisconsin and Minnesota on winter weekends. Iceboats are similar to sailboats, but they are fitted with runners, or “skates,” and are designed to run over the ice, rather than on water. Iceboats can move very quickly, and many ice boating enthusiasts are drawn to the sport because of the speed. Top iceboat racers can attain speeds up to five times the wind speed. If we know how fast an iceboat is moving, we can use integration to determine how far it travels. We revisit this question later in the chapter (see Example 1.27).

Determining distance from velocity is just one of many applications of integration. In fact, integrals are used in a wide variety of mechanical and physical applications. In this chapter, we first introduce the theory behind integration and use integrals to calculate areas. From there, we develop the Fundamental Theorem of Calculus, which relates differentiation and integration. We then study some basic integration techniques and briefly examine some applications.

## Learning Objectives

1.1.1. Use sigma (summation) notation to calculate sums and powers of integers.
1.1.2. Use the sum of rectangular areas to approximate the area under a curve.
1.1.3. Use Riemann sums to approximate area.

Archimedes was fascinated with calculating the areas of various shapes—in other words, the amount of space enclosed by the shape. He used a process that has come to be known as the *method of exhaustion*, which used smaller and smaller shapes, the areas of which could be calculated exactly, to fill an irregular region and thereby obtain closer and closer approximations to the total area. In this process, an area bounded by curves is filled with rectangles, triangles, and shapes with exact area formulas. These areas are then summed to approximate the area of the curved region.

In this section, we develop techniques to approximate the area between a curve, defined by a function $f(x),$ and the *x*-axis on a closed interval $[a, b].$ Like Archimedes, we first approximate the area under the curve using shapes of known area (namely, rectangles). By using smaller and smaller rectangles, we get closer and closer approximations to the area. Taking a limit allows us to calculate the exact area under the curve.

Let's start by introducing some notation to make the calculations easier. We then consider the case when $f(x)$ is continuous and nonnegative. Later in the chapter, we relax some of these restrictions and develop techniques that apply in more general cases.

## Sigma (Summation) Notation

As mentioned, we will use shapes of known area to approximate the area of an irregular region bounded by curves. This process often requires adding up long strings of numbers. To make it easier to write down these lengthy sums, we look at some new notation here, called **sigma notation** (also known as **summation notation**). The Greek capital letter $\Sigma,$ sigma, is used to express long sums of values in a compact form. For example, if we want to add all the integers from 1 to 20 without sigma notation, we have to write

$$1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + 13 + 14 + 15 + 16 + 17 + 18 + 19 + 20.$$

We could probably skip writing a couple of terms and write

$$1 + 2 + 3 + 4 + \cdots + 19 + 20,$$

which is better, but still cumbersome. With sigma notation, we write this sum as

$$\sum_{i=1}^{20} i,$$

which is much more compact.

Typically, sigma notation is presented in the form

$$\sum_{i=1}^{n} a_i$$

where $a_i$ describes the terms to be added, and the *i* is called the *index*. Each term is evaluated, then we sum all the values, beginning with the value when $i = 1$ and ending with the value when $i = n.$ For example, an expression like $\sum_{i=2}^{7} s_i$ is interpreted as $s_2 + s_3 + s_4 + s_5 + s_6 + s_7.$ Note that the index is used only to keep track of the terms to be added; it does not factor into the calculation of the sum itself. The index is therefore called a *dummy variable*. We can use any letter we like for the index. Typically, mathematicians use *i*, *j*, *k*, *m*, and *n* for indices.

Let's try a couple of examples of using sigma notation.

**EXAMPLE 1.1**

### Using Sigma Notation

a. Write in sigma notation and evaluate the sum of terms $3^i$ for $i = 1, 2, 3, 4, 5.$
b. Write the sum in sigma notation:

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25}.$$

[Show Solution]

**CHECKPOINT 1.1**

Write in sigma notation and evaluate the sum of terms $2^i$ for $i = 3, 4, 5, 6.$

The properties associated with the summation process are given in the following rule.

### RULE: PROPERTIES OF SIGMA NOTATION

Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ represent two sequences of terms and let *c* be a constant. The following properties hold for all positive integers *n* and for integers *m*, with $1 \leq m \leq n$.

1.

$$\sum_{i=1}^{n} c = nc \tag{1.1}$$

2.

$$\sum_{i=1}^{n} ca_i = c\sum_{i=1}^{n} a_i \tag{1.2}$$

3.

$$\sum_{i=1}^{n} \left(a_i + b_i\right) = \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i \tag{1.3}$$

4.

$$\sum_{i=1}^{n} \left(a_i - b_i\right) = \sum_{i=1}^{n} a_i - \sum_{i=1}^{n} b_i \tag{1.4}$$

5.

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{m} a_i + \sum_{i=m+1}^{n} a_i \tag{1.5}$$

## Proof

We prove properties 2. and 3. here, and leave proof of the other properties to the Exercises.

2. We have

$$\begin{aligned}\sum_{i=1}^{n} ca_i &= ca_1 + ca_2 + ca_3 + \cdots + ca_n \\ &= c\left(a_1 + a_2 + a_3 + \cdots + a_n\right) \\ &= c\sum_{i=1}^{n} a_i.\end{aligned}$$

3. We have

$$\begin{aligned}\sum_{i=1}^{n}\left(a_i + b_i\right) &= \left(a_1 + b_1\right) + \left(a_2 + b_2\right) + \left(a_3 + b_3\right) + \cdots + \left(a_n + b_n\right) \\ &= \left(a_1 + a_2 + a_3 + \cdots + a_n\right) + \left(b_1 + b_2 + b_3 + \cdots + b_n\right) \\ &= \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i.\end{aligned}$$

□

A few more formulas for frequently found functions simplify the summation process further. These are shown in the next rule, for **sums and powers of integers**, and we use them in the next set of examples.

### RULE: SUMS AND POWERS OF INTEGERS

1. The sum of *n* integers is given by

$$\sum_{i=1}^{n} i = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

2. The sum of consecutive integers squared is given by

$$\sum_{i=1}^{n} i^2 = 1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

3. The sum of consecutive integers cubed is given by

$$\sum_{i=1}^{n} i^3 = 1^3 + 2^3 + \cdots + n^3 = \frac{n^2(n+1)^2}{4}.$$

EXAMPLE 1.2

### Evaluation Using Sigma Notation

Write using sigma notation and evaluate:

a. The sum of the terms $(i-3)^2$ for $i = 1, 2, \ldots, 200$.
b. The sum of the terms $\left(i^3 - i^2\right)$ for $i = 1, 2, 3, 4, 5, 6$.

[Show Solution]

CHECKPOINT 1.2

Find the sum of the values of $4 + 3i$ for $i = 1, 2, \ldots, 100$.

EXAMPLE 1.3

### Finding the Sum of the Function Values

Find the sum of the values of $f(x) = x^3$ over the integers $1, 2, 3, \ldots, 10$.

[Show Solution]

CHECKPOINT 1.3

Evaluate the sum indicated by the notation $\sum_{k=1}^{20} (2k+1)$.

## Approximating Area

Now that we have the necessary notation, we return to the problem at hand: approximating the area under a curve. Let $f(x)$ be a continuous, nonnegative function defined on the closed interval $[a, b]$. We want to approximate the area *A* bounded by $f(x)$ above, the *x*-axis below, the line $x = a$ on the left, and the line $x = b$ on the right (Figure 1.2).

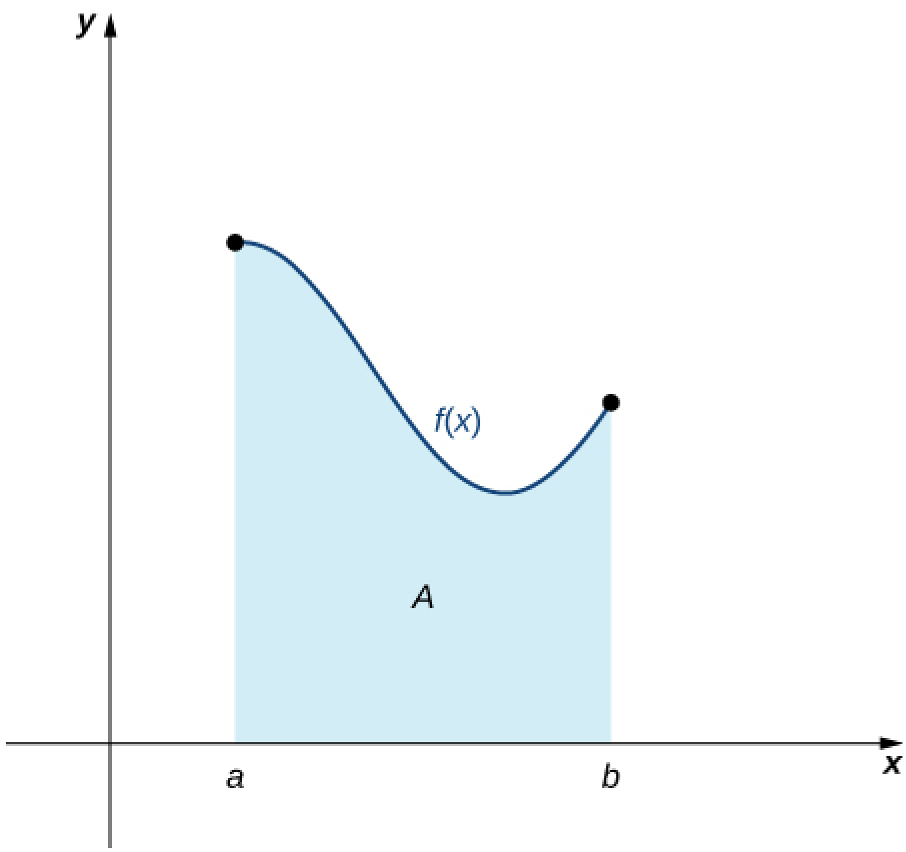



**Figure 1.2** An area (shaded region) bounded by the curve $f(x)$ at top, the *x*-axis at bottom, the line $x = a$ to the left, and the line $x = b$ at right.

How do we approximate the area under this curve? The approach is a geometric one. By dividing a region into many small shapes that have known area formulas, we can sum these areas and obtain a reasonable estimate of the true area. We begin by dividing the interval $[a, b]$ into *n* subintervals of equal width, $\frac{b-a}{n}$. We do this by selecting equally spaced points $x_0, x_1, x_2, \ldots, x_n$ with $x_0 = a, x_n = b,$ and

$$x_i - x_{i-1} = \frac{b - a}{n}$$

for $i = 1, 2, 3, \ldots, n.$

We denote the width of each subinterval with the notation Δ*x*, so $\Delta x = \frac{b-a}{n}$ and

$$x_i = x_0 + i\Delta x$$

for $i = 1, 2, 3, \ldots, n.$ This notion of dividing an interval $[a, b]$ into subintervals by selecting points from within the interval is used quite often in approximating the area under a curve, so let's define some relevant terminology.

**DEFINITION**

A set of points $P = \{x_i\}$ for $i = 0, 1, 2, \ldots, n$ with $a = x_0 < x_1 < x_2 < \cdots < x_n = b,$ which divides the interval $[a, b]$ into subintervals of the form $[x_0, x_1], [x_1, x_2], \ldots, [x_{n-1}, x_n]$ is called a **partition** of $[a, b].$ If the subintervals all have the same width, the set of points forms a **regular partition** of the interval $[a, b].$

We can use this regular partition as the basis of a method for estimating the area under the curve. We next examine two methods: the left-endpoint approximation and the right-endpoint approximation.

**RULE: LEFT-ENDPOINT APPROXIMATION**

On each subinterval $[x_{i-1}, x_i]$ (for $i = 1, 2, 3, \ldots, n$), construct a rectangle with width Δ*x* and height equal to $f(x_{i-1}),$ which is the function value at the left endpoint of the subinterval. Then the area of this rectangle is $f(x_{i-1})\Delta x.$ Adding the areas of all these rectangles, we get an approximate value for *A*

([Figure 1.3](#)). We use the notation $L_n$ to denote that this is a **left-endpoint approximation** of *A* using *n* subintervals.

$$\begin{aligned} A \approx L_n &= f(x_0)\Delta x + f(x_1)\Delta x + \cdots + f(x_{n-1})\Delta x \\ &= \sum_{i=1}^{n} f(x_{i-1})\Delta x \end{aligned} \quad (1.6)$$

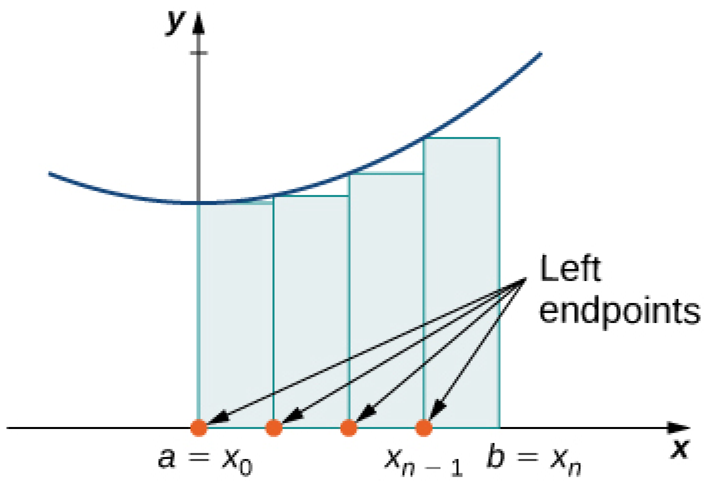



**Figure 1.3** In the left-endpoint approximation of area under a curve, the height of each rectangle is determined by the function value at the left of each subinterval.

The second method for approximating area under a curve is the right-endpoint approximation. It is almost the same as the left-endpoint approximation, but now the heights of the rectangles are determined by the function values at the right of each subinterval.

### RULE: RIGHT-ENDPOINT APPROXIMATION

Construct a rectangle on each subinterval $[x_{i-1}, x_i]$, only this time the height of the rectangle is determined by the function value $f(x_i)$ at the right endpoint of the subinterval. Then, the area of each rectangle is $f(x_i)\Delta x$ and the approximation for *A* is given by

1.7

$$\begin{aligned} A \approx R_n &= f(x_1)\Delta x + f(x_2)\Delta x + \cdots + f(x_n)\Delta x \\ &= \sum_{i=1}^{n} f(x_i)\Delta x. \end{aligned}$$

The notation $R_n$ indicates this is a **right-endpoint approximation** for *A* (Figure 1.4).

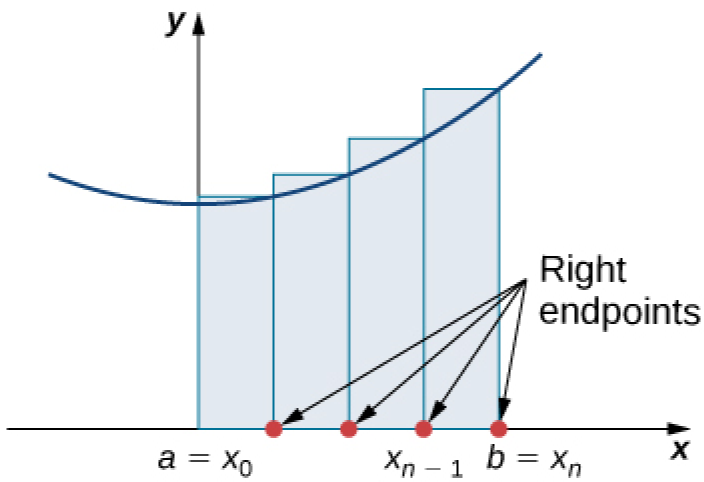



**Figure 1.4** In the right-endpoint approximation of area under a curve, the height of each rectangle is determined by the function value at the right of each subinterval. Note that the right-endpoint approximation differs from the left-endpoint approximation in Figure 1.3.

The graphs in Figure 1.5 represent the curve $f(x) = \frac{x^2}{2}$. In graph (a) we divide the region represented by the interval $[0, 3]$ into six subintervals, each of width 0.5. Thus, $\Delta x = 0.5$. We then form six rectangles by drawing vertical lines perpendicular to $x_{i-1}$, the left endpoint of each subinterval. We determine the height of each rectangle by calculating $f(x_{i-1})$ for $i = 1, 2, 3, 4, 5, 6$. The intervals are $[0, 0.5], [0.5, 1], [1, 1.5], [1.5, 2], [2, 2.5], [2.5, 3]$. We find the area of each rectangle by multiplying the height by the width. Then, the sum of the rectangular areas approximates the area between $f(x)$ and the *x*-axis. When the left endpoints are used to calculate height, we have a left-endpoint approximation. Thus,

$$\begin{aligned} A \approx L_6 &= \sum_{i=1}^{6} f(x_{i-1})\Delta x = f(x_0)\Delta x + f(x_1)\Delta x + f(x_2)\Delta x + f(x_3)\Delta x + f(x_4)\Delta x + f(x_5)\Delta x \\ &= f(0)0.5 + f(0.5)0.5 + f(1)0.5 + f(1.5)0.5 + f(2)0.5 + f(2.5)0.5 \\ &= (0)0.5 + (0.125)0.5 + (0.5)0.5 + (1.125)0.5 + (2)0.5 + (3.125)0.5 \\ &= 0 + 0.0625 + 0.25 + 0.5625 + 1 + 1.5625 \\ &= 3.4375. \end{aligned}$$

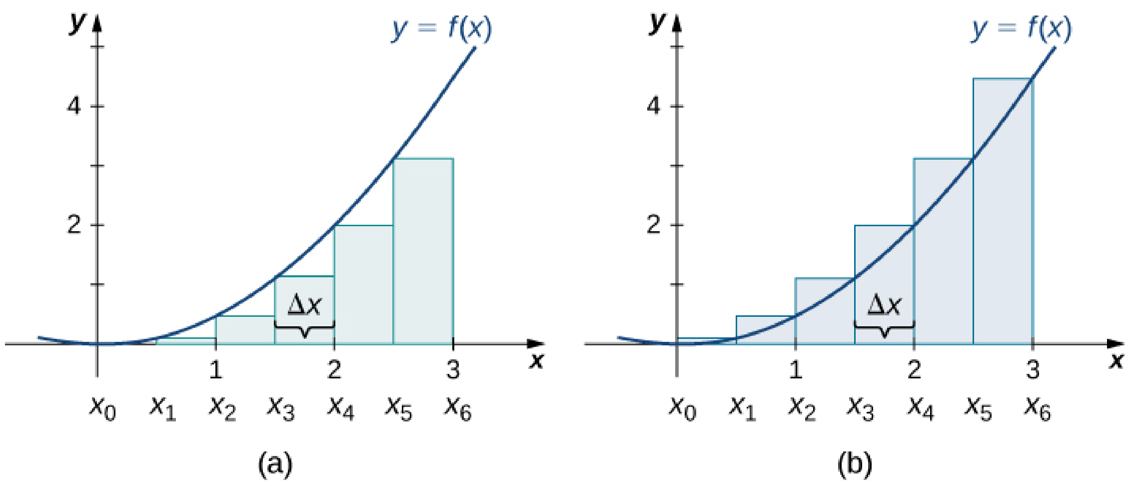



**Figure 1.5** Methods of approximating the area under a curve by using (a) the left endpoints and (b) the right endpoints.

In Figure 1.5(b), we draw vertical lines perpendicular to $x_i$ such that $x_i$ is the right endpoint of each subinterval, and calculate $f(x_i)$ for $i = 1, 2, 3, 4, 5, 6$. We multiply each $f(x_i)$ by $\Delta x$ to find the rectangular areas, and then add them. This is a right-endpoint approximation of the area under $f(x)$. Thus,

$$\begin{aligned} A \approx R_6 &= \sum_{i=1}^{6} f(x_i)\Delta x = f(x_1)\Delta x + f(x_2)\Delta x + f(x_3)\Delta x + f(x_4)\Delta x + f(x_5)\Delta x + f(x_6)\Delta x \\ &= f(0.5)0.5 + f(1)0.5 + f(1.5)0.5 + f(2)0.5 + f(2.5)0.5 + f(3)0.5 \\ &= (0.125)0.5 + (0.5)0.5 + (1.125)0.5 + (2)0.5 + (3.125)0.5 + (4.5)0.5 \\ &= 0.0625 + 0.25 + 0.5625 + 1 + 1.5625 + 2.25 \\ &= 5.6875. \end{aligned}$$

**EXAMPLE 1.4**

### Approximating the Area Under a Curve

Use both left-endpoint and right-endpoint approximations to approximate the area under the curve of $f(x) = x^2$ on the interval $[0, 2]$; use $n = 4$.

[Show Solution]

#### CHECKPOINT 1.4

Sketch left-endpoint and right-endpoint approximations for $f(x) = \frac{1}{x}$ on $[1, 2]$; use $n = 4$. Approximate the area using both methods.

Looking at Figure 1.5 and the graphs in Example 1.4, we can see that when we use a small number of intervals, neither the left-endpoint approximation nor the right-endpoint approximation is a particularly accurate estimate of the area under the curve. However, it seems logical that if we increase the number of points in our partition, our estimate of *A* will improve. We will have more rectangles, but each rectangle will be thinner, so we will be able to fit the rectangles to the curve more precisely.

We can demonstrate the improved approximation obtained through smaller intervals with an example. Let's explore the idea of increasing *n*, first in a left-endpoint approximation with four rectangles, then eight rectangles, and finally 32 rectangles. Then, let's do the same thing in a right-endpoint approximation, using the same sets of intervals, of the same curved region. Figure 1.8 shows the area of the region under the curve $f(x) = (x-1)^3 + 4$ on the interval $[0, 2]$ using a left-endpoint approximation where $n = 4$. The width of each rectangle is

$$\Delta x = \frac{2-0}{4} = \frac{1}{2}.$$

The area is approximated by the summed areas of the rectangles, or

$$\begin{aligned} L_4 &= f(0)(0.5) + f(0.5)(0.5) + f(1)(0.5) + f(1.5)0.5 \\ &= 7.5. \end{aligned}$$

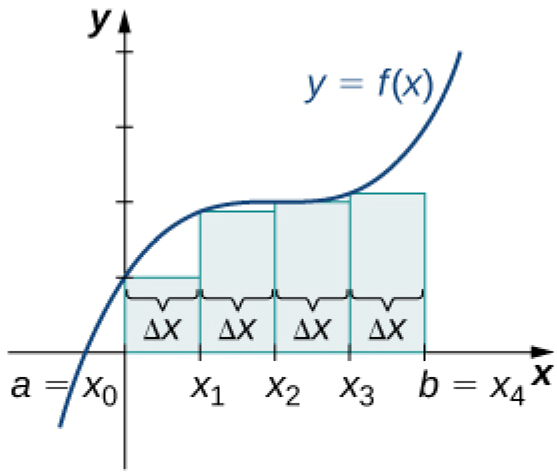



**Figure 1.8** With a left-endpoint approximation and dividing the region from *a* to *b* into four equal intervals, the area under the curve is approximately equal to the sum of the areas of the rectangles.

Figure 1.9 shows the same curve divided into eight subintervals. Comparing the graph with four rectangles in Figure 1.8 with this graph with eight rectangles, we can see there appears to be less white space under the curve when $n = 8.$ This white space is area under the curve we are unable to include using our approximation. The area of the rectangles is

$$\begin{aligned} L_8 &= f(0)(0.25) + f(0.25)(0.25) + f(0.5)(0.25) + f(0.75)(0.25) \\ &\quad + f(1)(0.25) + f(1.25)(0.25) + f(1.5)(0.25) + f(1.75)(0.25) \\ &= 7.75. \end{aligned}$$

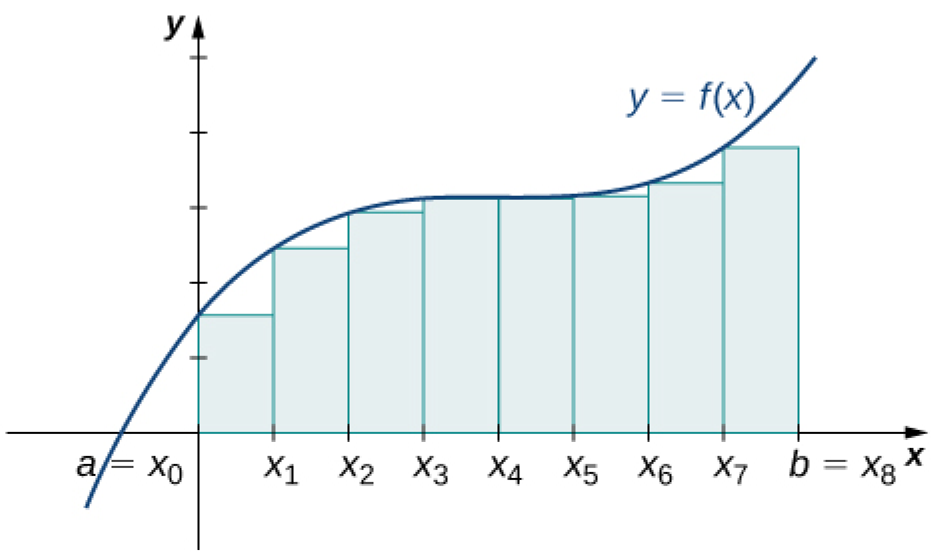



**Figure 1.9** The region under the curve is divided into $n = 8$ rectangular areas of equal width

for a left-endpoint approximation.

The graph in Figure 1.10 shows the same function with 32 rectangles inscribed under the curve. There appears to be little white space left. The area occupied by the rectangles is

$$\begin{aligned} L_{32} &= f(0)(0.0625) + f(0.0625)(0.0625) + f(0.125)(0.0625) + \cdots + f(1.9375)(0.0625) \\ &= 7.9375. \end{aligned}$$

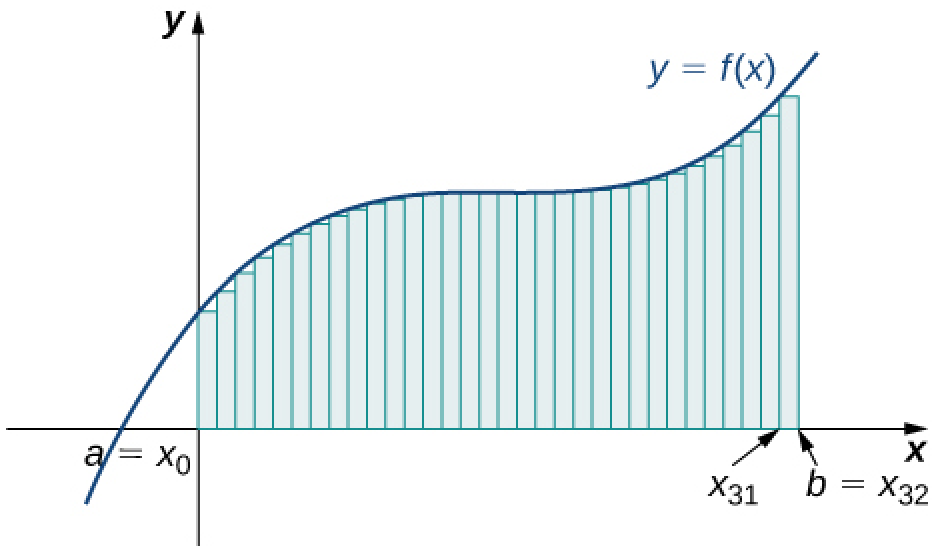



**Figure 1.10** Here, 32 rectangles are inscribed under the curve for a left-endpoint approximation.

We can carry out a similar process for the right-endpoint approximation method. A right-endpoint approximation of the same curve, using four rectangles (Figure 1.11), yields an area

$$\begin{aligned} R_4 &= f(0.5)(0.5) + f(1)(0.5) + f(1.5)(0.5) + f(2)(0.5) \\ &= 8.5. \end{aligned}$$

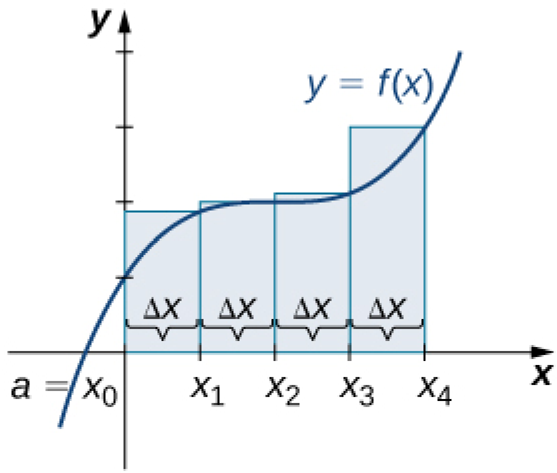



**Figure 1.11** Now we divide the area under the curve into four equal subintervals for a right-endpoint approximation.

Dividing the region over the interval $[0, 2]$ into eight rectangles results in $\Delta x = \frac{2-0}{8} = 0.25.$ The graph is shown in Figure 1.12. The area is

$$\begin{aligned} R_8 &= f(0.25)(0.25) + f(0.5)(0.25) + f(0.75)(0.25) + f(1)(0.25) \\ &+ f(1.25)(0.25) + f(1.5)(0.25) + f(1.75)(0.25) + f(2)(0.25) \\ &= 8.25. \end{aligned}$$

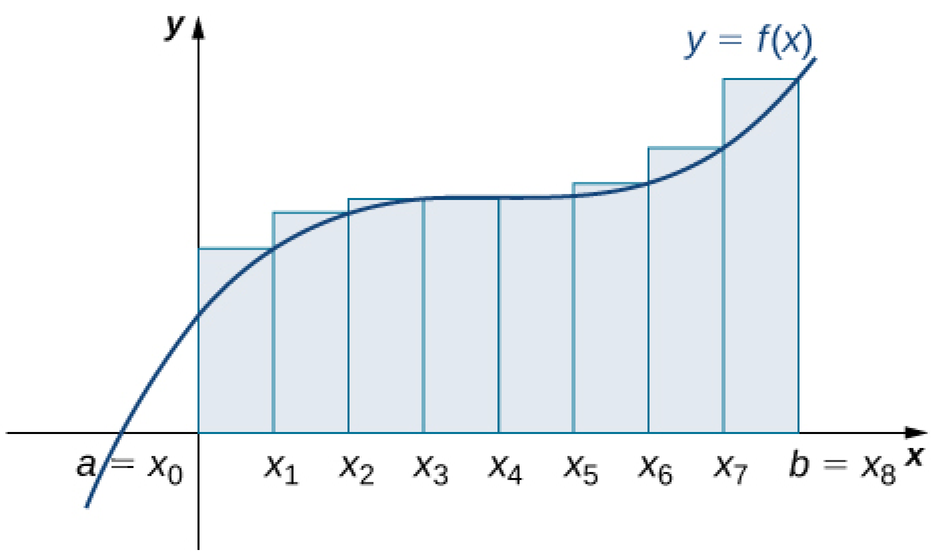



**Figure 1.12** Here we use right-endpoint approximation for a region divided into eight equal subintervals.

Last, the right-endpoint approximation with $n = 32$ is close to the actual area (Figure 1.13). The area is approximately

$$\begin{aligned} R_{32} &= f(0.0625)(0.0625) + f(0.125)(0.0625) + f(0.1875)(0.0625) + \cdots + f(2)(0.0625) \\ &= 8.0625. \end{aligned}$$

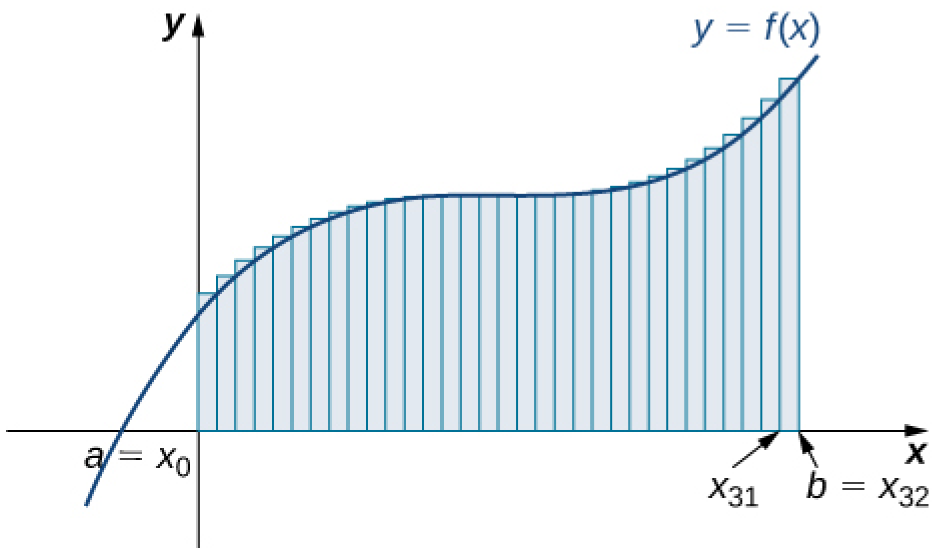



**Figure 1.13** The region is divided into 32 equal subintervals for a right-endpoint approximation.

Based on these figures and calculations, it appears we are on the right track; the rectangles appear to approximate the area under the curve better as *n* gets larger. Furthermore, as *n* increases, both the left-endpoint and right-endpoint approximations appear to approach an area of 8 square units. Table 1.1 shows a numerical comparison of the left- and right-endpoint methods. The idea that the approximations of the area under the curve get better and better as *n* gets larger and larger is very important, and we now explore this idea in more detail.

| Values of $n$ | Approximate Area $L_n$ | Approximate Area $R_n$ |
|---|---|---|
| $n = 4$ | 7.5 | 8.5 |
| $n = 8$ | 7.75 | 8.25 |
| $n = 32$ | 7.94 | 8.06 |

**Table 1.1** Converging Values of Left- and Right-Endpoint Approximations as *n* Increases

## Forming Riemann Sums

So far we have been using rectangles to approximate the area under a curve. The heights of these rectangles have been determined by evaluating the function at either the right or left endpoints of the subinterval $\left[x_{i-1}, x_i\right]$. In reality, there is no reason to restrict evaluation of the function to one of these two points only. We could evaluate the function at any point $x_i$ in the subinterval $\left[x_{i-1}, x_i\right]$, and use $f\left(x_i^*\right)$ as the height of our rectangle. This gives us an estimate for the area of the form

$$A \approx \sum_{i=1}^{n} f\left(x_i^*\right)\Delta x.$$

A sum of this form is called a Riemann sum, named for the 19th-century mathematician Bernhard Riemann, who developed the idea.

**DEFINITION**

Let $f(x)$ be defined on a closed interval $[a, b]$ and let $P$ be a regular partition of $[a, b]$. Let $\Delta x$ be the width of each subinterval $\left[x_{i-1}, x_i\right]$ and for each $i$, let $x_i^*$ be any point in $\left[x_{i-1}, x_i\right]$. A **Riemann sum** is defined for $f(x)$ as

$$\sum_{i=1}^{n} f\left(x_i^*\right)\Delta x.$$

Recall that with the left- and right-endpoint approximations, the estimates seem to get better and better as *n* get larger and larger. The same thing happens with Riemann sums. Riemann sums give better approximations for larger values of *n*. We are now ready to define the area under a curve in terms of Riemann sums.

**DEFINITION**

Let $f(x)$ be a continuous, nonnegative function on an interval $[a, b]$, and let $\sum_{i=1}^{n} f\left(x_i^*\right)\Delta x$ be a Riemann sum for $f(x)$. Then, the **area under the curve** $y = f(x)$

on $[a, b]$ is given by

$$A = \lim_{n \to \infty} \sum_{i=1}^{n} f\left(x_i^*\right) \Delta x.$$

**MEDIA**

See a graphical demonstration of the construction of a Riemann sum.

Some subtleties here are worth discussing. First, note that taking the limit of a sum is a little different from taking the limit of a function $f(x)$ as $x$ goes to infinity. Limits of sums are discussed in detail in the chapter on Sequences and Series; however, for now we can assume that the computational techniques we used to compute limits of functions can also be used to calculate limits of sums.

Second, we must consider what to do if the expression converges to different limits for different choices of $\left\{x_i^*\right\}$. Fortunately, this does not happen. Although the proof is beyond the scope of this text, it can be shown that if $f(x)$ is continuous on the closed interval $[a, b]$, then $\lim_{n \to \infty} \sum_{i=1}^{n} f\left(x_i^*\right) \Delta x$ exists and is unique (in other words, it does not depend on the choice of $\left\{x_i^*\right\}$).

We look at some examples shortly. But, before we do, let's take a moment and talk about some specific choices for $\left\{x_i^*\right\}$. Although any choice for $\left\{x_i^*\right\}$ gives us an estimate of the area under the curve, we don't necessarily know whether that estimate is too high (overestimate) or too low (underestimate). If it is important to know whether our estimate is high or low, we can select our value for $\left\{x_i^*\right\}$ to guarantee one result or the other.

If we want an overestimate, for example, we can choose $\left\{x_i^*\right\}$ such that for $i = 1, 2, 3, \ldots, n$, $f\left(x_i^*\right) \ge f(x)$ for all $x \in \left[x_{i-1}, x_i\right]$. In other words, we choose $\left\{x_i^*\right\}$ so that for $i = 1, 2, 3, \ldots, n$, $f\left(x_i^*\right)$ is the maximum function value on the interval $\left[x_{i-1}, x_i\right]$. If we select $\left\{x_i^*\right\}$ in this way, then the Riemann sum $\sum_{i=1}^{n} f\left(x_i^*\right) \Delta x$ is called an **upper sum**.

Similarly, if we want an underestimate, we can choose $\{x_i^*\}$ so that for $i = 1, 2, 3,\ldots, n, f\left(x_i^*\right)$ is the minimum function value on the interval $\left[x_{i-1}, x_i\right]$. In this case, the associated Riemann sum is called a **lower sum**. Note that if $f(x)$ is either increasing or decreasing throughout the interval $[a, b]$, then the maximum and minimum values of the function occur at the endpoints of the subintervals, so the upper and lower sums are just the same as the left- and right-endpoint approximations.

**EXAMPLE 1.5**

### Finding Lower and Upper Sums

Find a lower sum for $f(x) = 10 - x^2$ on $[1, 2]$; let $n = 4$ subintervals.

[Show Solution]

**CHECKPOINT 1.5**

a. Find an upper sum for $f(x) = 10 - x^2$ on $[1, 2]$; let $n = 4$.
b. Sketch the approximation.

**EXAMPLE 1.6**

### Finding Lower and Upper Sums for $f(x) = \sin x$

Find a lower sum for $f(x) = \sin x$ over the interval $[a, b] = \left[0, \frac{\pi}{2}\right]$; let $n = 6$.

[Show Solution]

## CHECKPOINT 1.6

Using the function $f(x) = \sin x$ over the interval $\left[0, \frac{\pi}{2}\right]$, find an upper sum; let $n = 6$.

# Section 1.1 Exercises

1. State whether the given sums are equal or unequal.

   a. $\sum_{i=1}^{10} i$ and $\sum_{k=1}^{10} k$

   b. $\sum_{i=1}^{10} i$ and $\sum_{i=6}^{15} (i - 5)$

   c. $\sum_{i=1}^{10} i(i - 1)$ and $\sum_{j=0}^{9} (j + 1)j$

   d. $\sum_{i=1}^{10} i(i - 1)$ and $\sum_{k=1}^{10} \left(k^2 - k\right)$

In the following exercises, use the rules for sums of powers of integers to compute the sums.

2. $\sum_{i=5}^{10} i$

3. $\sum_{i=5}^{10} i^2$

Suppose that $\sum_{i=1}^{100} a_i = 15$ and $\sum_{i=1}^{100} b_i = -12$. In the following exercises, compute the sums.

4. $\sum_{i=1}^{100} \left(a_i + b_i\right)$

5. $\sum_{i=1}^{100} \left(a_i - b_i\right)$

6. $\sum_{i=1}^{100} \left(3a_i - 4b_i\right)$

7. $\sum_{i=1}^{100} \left(5a_i + 4b_i\right)$

In the following exercises, use summation properties and formulas to rewrite and evaluate the sums.

8. $\sum_{k=1}^{20} 100\left(k^2 - 5k + 1\right)$

9. $\sum_{j=1}^{50} \left(j^2 - 2j\right)$

10. $\sum_{j=11}^{20} \left(j^2 - 10j\right)$

11. $\sum_{k=1}^{25} \left[(2k)^2 - 100k\right]$

Let $L_n$ denote the left-endpoint sum using *n* subintervals and let $R_n$ denote the corresponding right-endpoint sum. In the following exercises, compute the indicated left and right sums for the given functions on the indicated interval.

12. $L_4$ for $f(x) = \frac{1}{x-1}$ on $[2, 3]$

13. $R_4$ for $g(x) = \cos(\pi x)$ on $[0, 1]$

14. $L_6$ for $f(x) = \frac{1}{x(x-1)}$ on $[2, 5]$

15. $R_6$ for $f(x) = \frac{1}{x(x-1)}$ on $[2, 5]$

16. $R_4$ for $\frac{1}{x^2+1}$ on $[-2, 2]$

17. $L_4$ for $\frac{1}{x^2+1}$ on $[-2, 2]$

18. $R_4$ for $x^2 - 2x + 1$ on $[0, 2]$

19. $L_8$ for $x^2 - 2x + 1$ on $[0, 2]$

20. Compute the left and right Riemann sums—$L_4$ and $R_4$, respectively—for $f(x) = (2 - |x|)$ on $[-2, 2]$. Compute their average value and compare it with the area under the graph of $f$.

21. Compute the left and right Riemann sums—$L_6$ and $R_6$, respectively—for $f(x) = (3 - |3 - x|)$ on $[0, 6]$. Compute their average value and compare it with the area under the graph of $f$.

22. Compute the left and right Riemann sums—$L_4$ and $R_4$, respectively—for $f(x) = \sqrt{4 - x^2}$ on $[-2, 2]$ and compare their values.

23. Compute the left and right Riemann sums—$L_6$ and $R_6$, respectively—for $f(x) = \sqrt{9 - (x-3)^2}$ on $[0, 6]$ and compare their values.

Express the following endpoint sums in sigma notation but do not evaluate them.

24. $L_{30}$ for $f(x) = x^2$ on $[1, 2]$

25. $L_{10}$ for $f(x) = \sqrt{4 - x^2}$ on $[-2, 2]$

26. $R_{20}$ for $f(x) = \sin x$ on $[0, \pi]$

27. $R_{100}$ for $\ln x$ on $[1, e]$

In the following exercises, graph the function then use a calculator or a computer program to evaluate the following left and right endpoint sums. Is the area under the curve on the given interval better approximated by the left Riemann sum or right Riemann sum? If the two agree, say "neither."

28. **[T]** $L_{100}$ and $R_{100}$ for $y = x^2 - 3x + 1$ on the interval $[-1, 1]$

29. **[T]** $L_{100}$ and $R_{100}$ for $y = x^2$ on the interval $[0, 1]$

30. **[T]** $L_{50}$ and $R_{50}$ for $y = \frac{x+1}{x^2-1}$ on the interval $[2, 4]$

31. **[T]** $L_{100}$ and $R_{100}$ for $y = x^3$ on the interval $[-1, 1]$

32. **[T]** $L_{50}$ and $R_{50}$ for $y = \tan(x)$ on the interval $\left[0, \frac{\pi}{4}\right]$

33. **[T]** $L_{100}$ and $R_{100}$ for $y = e^{2x}$ on the interval $[-1, 1]$

34. Let $t_j$ denote the time that it took Tejay van Garteren to ride the $j$th stage of the Tour de France in 2014. If there were a total of 21 stages, interpret $\sum_{j=1}^{21} t_j$.

35. Let $r_j$ denote the total rainfall in Portland on the $j$th day of the year in 2009. Interpret $\sum_{j=1}^{31} r_j$.

36. Let $d_j$ denote the hours of daylight and $\delta_j$ denote the increase in the hours of daylight from day $j-1$ to day $j$ in Fargo, North Dakota, on the $j$th day of the year. Interpret $d_1 + \sum_{j=2}^{365} \delta_j$.

37. To help get in shape, Joe gets a new pair of running shoes. If Joe runs 1 mi each day in week 1 and adds $\frac{1}{10}$ mi to his daily routine each week, what is the total mileage on Joe's shoes after 25 weeks?

38. The following table gives approximate values of the average annual atmospheric rate of increase in carbon dioxide ($CO_2$) each decade since 1960, in parts per million (ppm). Estimate the total increase in atmospheric $CO_2$ between 1964 and 2013.

| **Decade** | **Ppm/y** |
|---|---|
| 1964–1973 | 1.07 |
| 1974–1983 | 1.34 |
| 1984–1993 | 1.40 |
| 1994–2003 | 1.87 |
| 2004–2013 | 2.07 |

**Table 1.2** Average Annual Atmospheric $CO_2$ Increase, 1964–2013 *Source*: http://www.esrl.noaa.gov/gmd/ccgg/trends/.

39. The following table gives the approximate increase in sea level in inches over 20 years starting in the given year. Estimate the net change in mean sea level from 1870 to 2010.

| Starting Year | 20-Year Change |
|---|---|
| 1870 | 0.3 |
| 1890 | 1.5 |
| 1910 | 0.2 |
| 1930 | 2.8 |
| 1950 | 0.7 |
| 1970 | 1.1 |
| 1990 | 1.5 |

**Table 1.3** Approximate 20-Year Sea Level Increases, 1870–1990 *Source*: http://link.springer.com/article/10.1007%2Fs10712-011-9119-1

40. The following table gives the approximate increase in dollars in the average price of a gallon of gas per decade since 1950. If the average price of a gallon of gas in 2010 was $2.60, what was the average price of a gallon of gas in 1950?

| Starting Year | 10-Year Change |
|---|---|
| 1950 | 0.03 |
| 1960 | 0.05 |
| 1970 | 0.86 |
| 1980 | –0.03 |
| 1990 | 0.29 |
| 2000 | 1.12 |

**Table 1.4** Approximate 10-Year Gas Price Increases, 1950–2000 *Source*: http://epb.lbl.gov/homepages/Rick_Diamond/docs/lbnl55011-trends.pdf.

41. The following table gives the percent growth of the U.S. population beginning in July of the year indicated. If the U.S. population was 281,421,906 in July 2000, estimate the U.S. population in July 2010.

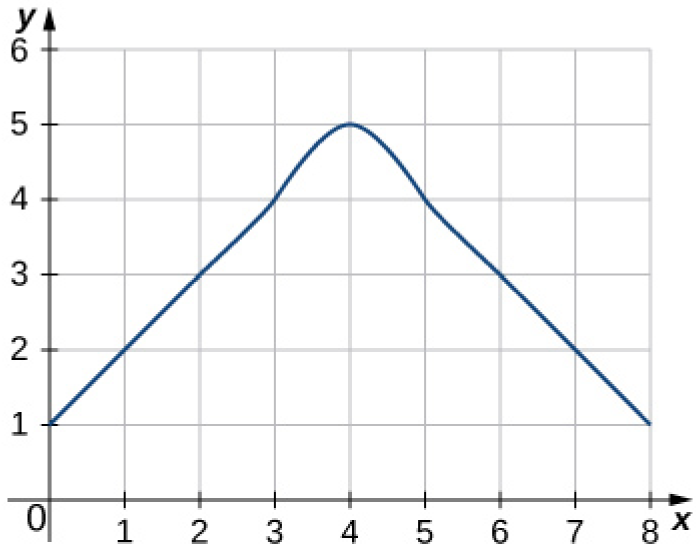
| Year | % Change/Year |
|---|---|
| 2000 | 1.12 |
| 2001 | 0.99 |
| 2002 | 0.93 |
| 2003 | 0.86 |
| 2004 | 0.93 |
| 2005 | 0.93 |
| 2006 | 0.97 |
| 2007 | 0.96 |
| 2008 | 0.95 |
| 2009 | 0.88 |

**Table 1.5** Annual Percentage Growth of U.S. Population, 2000–2009 *Source*: http://www.census.gov/popest/data.

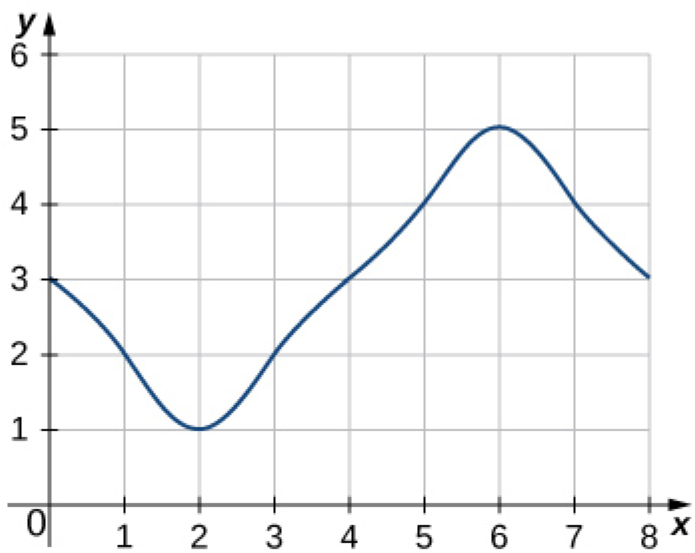
(*Hint:* To obtain the population in July 2001, multiply the population in July 2000 by 1.0112 to get 284,573,831.)

In the following exercises, estimate the areas under the curves by computing the left Riemann sums, $L_8$.

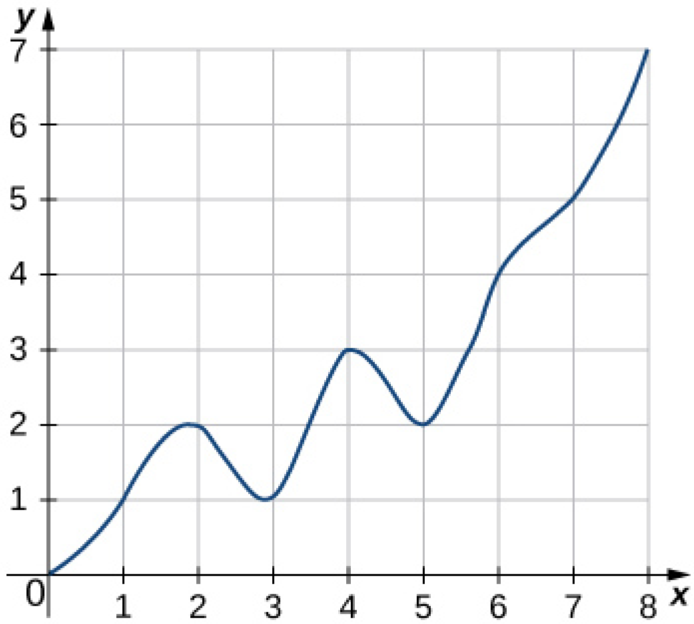
42.



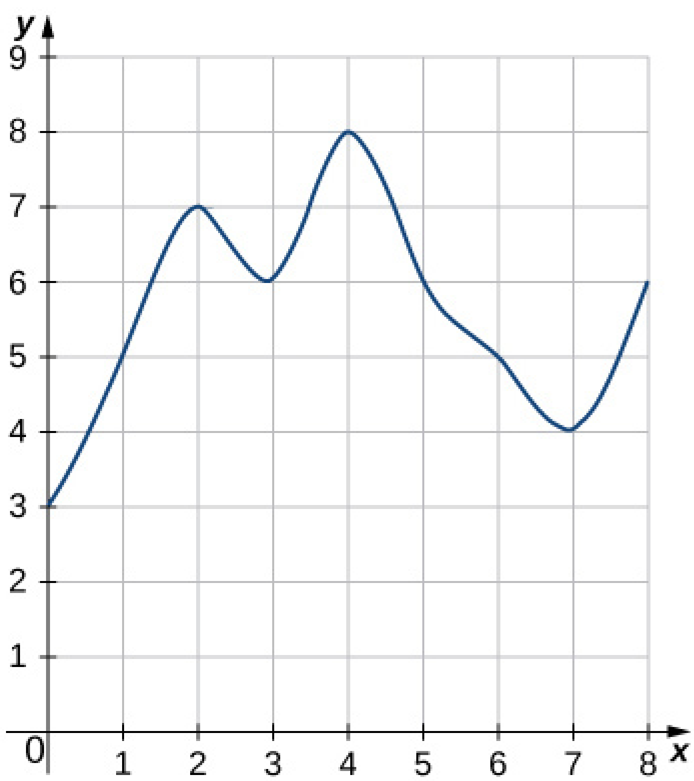

43.




44.




45.




46. **[T]** Use a computer algebra system to compute the Riemann sum, $L_N$, for $N = 10, 30, 50$ for $f(x) = \sqrt{1 - x^2}$ on $[-1, 1]$.

47. **[T]** Use a computer algebra system to compute the Riemann sum, $L_N$, for $N = 10, 30, 50$ for $f(x) = \frac{1}{\sqrt{1+x^2}}$ on $[-1, 1]$.

48. **[T]** Use a computer algebra system to compute the Riemann sum, $L_N$, for $N = 10, 30, 50$ for $f(x) = \sin^2 x$ on $[0, 2\pi]$. Compare these estimates with $\pi$.

In the following exercises, use a calculator or a computer program to evaluate the endpoint sums $R_N$ and $L_N$ for $N = 1,10,100$. How do these estimates compare with the exact answers, which you can find via geometry?

49. **[T]** $y = \cos(\pi x)$ on the interval $[0, 1]$

50. **[T]** $y = 3x + 2$ on the interval $[3, 5]$

In the following exercises, use a calculator or a computer program to evaluate the endpoint sums $R_N$ and $L_N$ for $N = 1,10,100$.

51. **[T]** $y = x^4 - 5x^2 + 4$ on the interval $[-2, 2]$, which has an exact area of $\frac{32}{15}$

52. **[T]** $y = \ln x$ on the interval $[1, 2]$, which has an exact area of $2\ln(2) - 1$

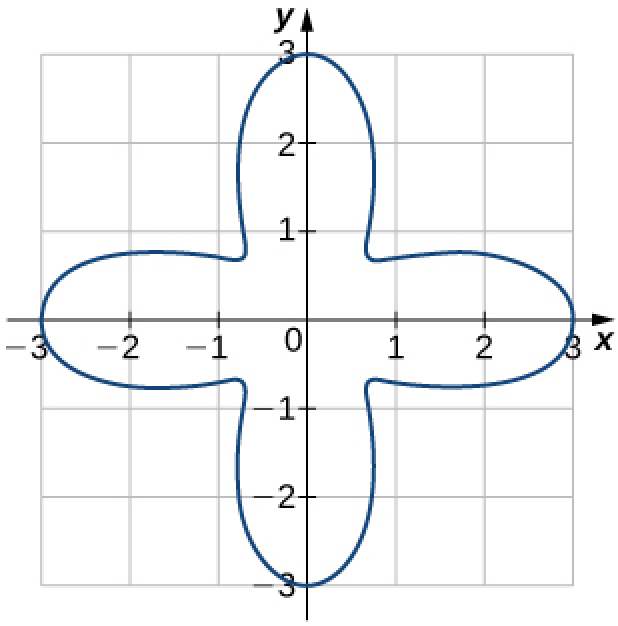
53. Explain why, if $f(a) \geq 0$ and *f* is increasing on $[a, b]$, that the left endpoint estimate is a lower bound for the area below the graph of *f* on $[a, b]$.

54. Explain why, if $f(b) \geq 0$ and *f* is decreasing on $[a, b]$, that the left endpoint estimate is an upper bound for the area below the graph of *f* on $[a, b]$.

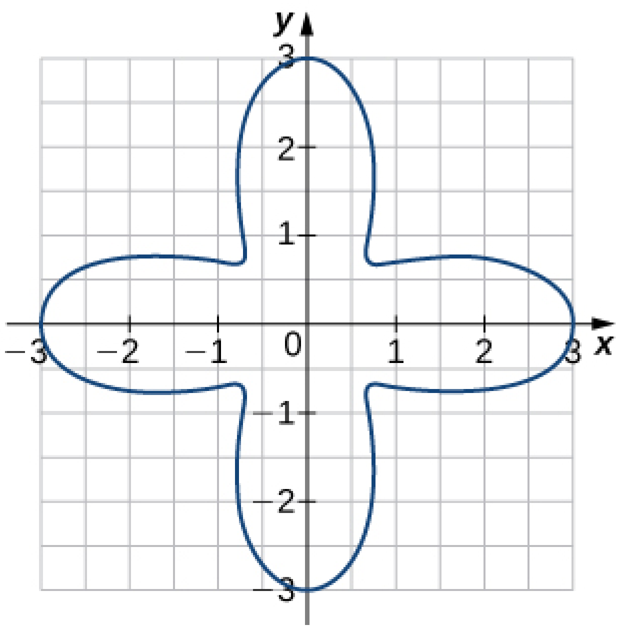
55. Show that, in general, $R_N - L_N = (b - a) \times \frac{f(b) - f(a)}{N}$.

56. Explain why, if *f* is increasing on $[a, b]$, the error between either $L_N$ or $R_N$ and the area *A* below the graph of *f* is at most $(b - a)\frac{f(b) - f(a)}{N}$.

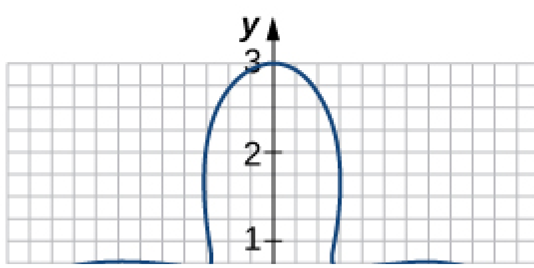
57. For each of the three graphs:
   a. Obtain a lower bound $L(A)$ for the area enclosed by the curve by adding the areas of the squares *enclosed completely* by the curve.
   b. Obtain an upper bound $U(A)$ for the area by adding to $L(A)$ the areas $B(A)$ of the squares *enclosed partially* by the curve.

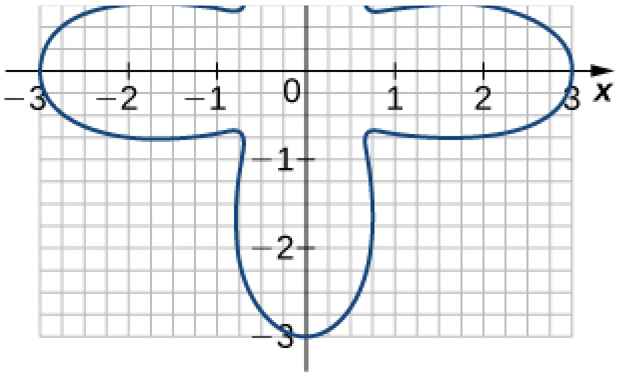



Graph 1




Graph 2

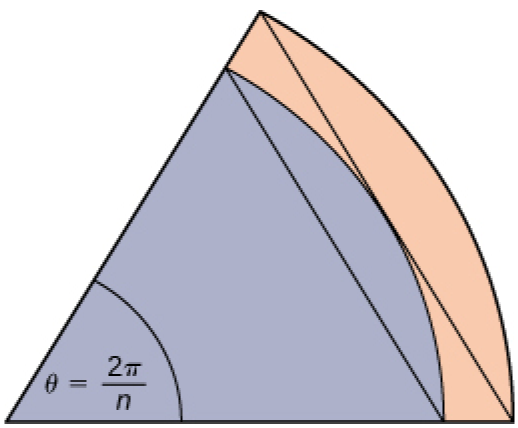





Graph 3

58. In the previous exercise, explain why $L(A)$ gets no smaller while $U(A)$ gets no larger as the squares are subdivided into four boxes of equal area.

59. A unit circle is made up of *n* wedges equivalent to the inner wedge in the figure. The base of the inner triangle is 1 unit and its height is $\sin\left(\frac{\pi}{n}\right)$. The base of the outer triangle is $B = \cos\left(\frac{\pi}{n}\right) + \sin\left(\frac{\pi}{n}\right)\tan\left(\frac{\pi}{n}\right)$ and the height is $H = B\sin\left(\frac{2\pi}{n}\right)$. Use this information to argue that the area of a unit circle is equal to $\pi$.



# Learning Objectives

1.2.1. State the definition of the definite integral.
1.2.2. Explain the terms integrand, limits of integration, and variable of integration.
1.2.3. Explain when a function is integrable.
1.2.4. Describe the relationship between the definite integral and net area.
1.2.5. Use geometry and the properties of definite integrals to evaluate them.
1.2.6. Calculate the average value of a function.

In the preceding section we defined the area under a curve in terms of Riemann sums:

$$A = \lim_{n \to \infty} \sum_{i=1}^{n} f\left(x_i^*\right)\Delta x.$$

However, this definition came with restrictions. We required $f(x)$ to be continuous and nonnegative. Unfortunately, real-world problems don't always meet these restrictions. In this section, we look at how to apply the concept of the area under the curve to a broader set of functions through the use of the definite integral.

# Definition and Notation

The definite integral generalizes the concept of the area under a curve. We lift the requirements that $f(x)$ be continuous and nonnegative, and define the definite integral as follows.

**DEFINITION**

If $f(x)$ is a function defined on an interval $[a, b]$, the **definite integral** of *f* from *a* to *b* is given by

$$\int_a^b f(x)dx = \lim_{n \to \infty} \sum_{i=1}^{n} f\left(x_i^*\right)\Delta x, \tag{1.8}$$

provided the limit exists. If this limit exists, the function $f(x)$ is said to be integrable on $[a, b]$, or is an **integrable function**.

The integral symbol in the previous definition should look familiar. We have seen similar notation in the chapter on Applications of Derivatives, where we used the indefinite integral symbol (without the *a* and *b* above and below) to represent an antiderivative. Although the notation for indefinite integrals may look similar to the notation for a definite integral, they are not the same. A definite integral is a number. An indefinite integral is a family of functions. Later in this chapter we examine how these concepts are related. However, close attention should always be paid to notation so we know whether we're working with a definite integral or an indefinite integral.

Integral notation goes back to the late seventeenth century and is one of the contributions of Gottfried Wilhelm Leibniz, who is often considered to be the codiscoverer of calculus, along with Isaac Newton. The integration symbol $\int$ is an elongated S, suggesting sigma or summation. On a definite integral, above and below the summation symbol are the boundaries of the interval, $[a, b]$. The numbers *a* and *b* are *x*-values and are called the **limits of integration**; specifically, *a* is the lower limit and *b* is the upper limit. To clarify, we are using the word *limit* in two different ways in the context of the definite integral. First, we talk about the limit of a sum as $n \to \infty$. Second, the boundaries of the region are called the *limits of integration*.

We call the function $f(x)$ the **integrand**, and the *dx* indicates that $f(x)$ is a function with respect to *x*, called the **variable of integration.** Note that, like the index in a sum, the variable of integration is a dummy variable, and has no impact on the computation of the integral. We could use any variable we like as the variable of integration:

$$\int_a^b f(x)dx = \int_a^b f(t)dt = \int_a^b f(u)du$$

Previously, we discussed the fact that if $f(x)$ is continuous on $[a, b]$, then the limit $\lim\limits_{n \to \infty} \sum\limits_{i=1}^{n} f\left(x_i^*\right)\Delta x$ exists and is unique. This leads to the following theorem, which we state without proof.

**THEOREM 1.1**

### Continuous Functions Are Integrable

If $f(x)$ is continuous on $[a, b]$, then *f* is integrable on $[a, b]$.

Functions that are not continuous on $[a, b]$ may still be integrable, depending on the nature of the discontinuities. For example, functions with a finite number of jump discontinuities on a closed interval are integrable.

It is also worth noting here that we have retained the use of a regular partition in the Riemann sums. This restriction is not strictly necessary. Any partition can be used to form a Riemann sum. However, if a nonregular partition is used to define the definite integral, it is not sufficient to take the limit as the number of subintervals goes to infinity. Instead, we must take the limit as the width of the largest subinterval goes to zero. This introduces a little more complex notation in our limits and makes the calculations more difficult without really gaining much additional insight, so we stick with regular partitions for the Riemann sums.

**EXAMPLE 1.7**

### Evaluating an Integral Using the Definition

Use the definition of the definite integral to evaluate $\int_0^2 x^2 dx.$ Use a right-endpoint approximation to generate the Riemann sum.

[Show Solution]

**CHECKPOINT 1.7**

Use the definition of the definite integral to evaluate $\int_0^3 (2x - 1)dx.$ Use a right-endpoint approximation to generate the Riemann sum.

## Evaluating Definite Integrals

Evaluating definite integrals this way can be quite tedious because of the complexity of the calculations. Later in this chapter we develop techniques for evaluating definite integrals *without* taking limits of Riemann sums. However, for now, we can rely on the fact that definite integrals represent the area under the curve, and we can evaluate definite integrals by using geometric formulas to calculate that area. We do this to confirm that definite integrals do, indeed, represent areas, so we can then discuss what to do in the case of a curve of a function dropping below the *x*-axis.

**EXAMPLE 1.8**

### Using Geometric Formulas to Calculate Definite Integrals

Use the formula for the area of a circle to evaluate $\int_3^6 \sqrt{9-(x-3)^2}\,dx.$

**[Show Solution]**

**CHECKPOINT 1.8**

Use the formula for the area of a trapezoid to evaluate $\int_2^4 (2x+3)\,dx.$

## Area and the Definite Integral

When we defined the definite integral, we lifted the requirement that $f(x)$ be nonnegative. But how do we interpret "the area under the curve" when $f(x)$ is negative?

### Net Signed Area

Let us return to the Riemann sum. Consider, for example, the function $f(x) = 2 - 2x^2$ (shown in Figure 1.17) on the interval $[0, 2]$. Use $n = 8$ and choose $\{x_i^*\}$ as the left endpoint of each interval. Construct a rectangle on each subinterval of height $f\left(x_i^*\right)$ and width $\Delta x$. When $f\left(x_i^*\right)$ is positive, the product $f\left(x_i^*\right)\Delta x$ represents the area of the rectangle, as before. When $f\left(x_i^*\right)$ is negative, however, the product $f\left(x_i^*\right)\Delta x$ represents the *negative* of the area of the rectangle. The Riemann sum then becomes

$$\sum_{i=1}^{8} f\left(x_i^*\right)\Delta x = (\text{Area of rectangles above the } x\text{-axis}) - (\text{Area of rectangles below the } x\text{-axis})$$

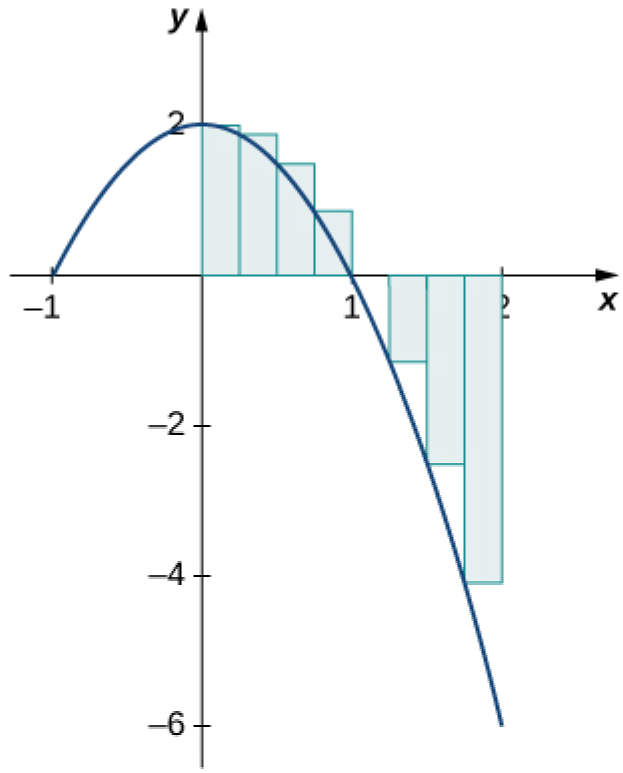



**Figure 1.17** For a function that is partly negative, the Riemann sum is the area of the rectangles above the *x*-axis less the area of the rectangles below the *x*-axis.

Taking the limit as $n \to \infty$, the Riemann sum approaches the area between the curve above the *x*-axis and the *x*-axis, less the area between the curve below the *x*-axis and the *x*-axis, as shown in Figure 1.18. Then,

$$\begin{aligned} \int_0^2 f(x)dx &= \lim_{n \to \infty} \sum_{i=1}^{n} f\left(c_i\right)\Delta x \\ &= A_1 - A_2. \end{aligned}$$

The quantity $A_1 - A_2$ is called the **net signed area.**

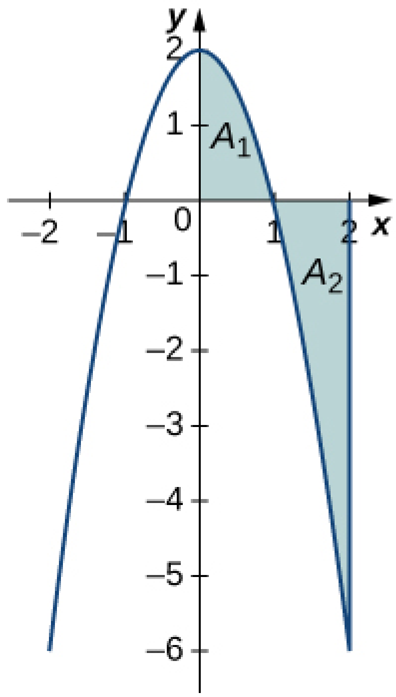



**Figure 1.18** In the limit, the definite integral equals area $A_1$ less area $A_2$, or the net signed area.

Notice that net signed area can be positive, negative, or zero. If the area above the *x*-axis is larger, the net signed area is positive. If the area below the *x*-axis is larger, the net signed area is negative. If the areas above and below the *x*-axis are equal, the net signed area is zero.

**EXAMPLE 1.9**

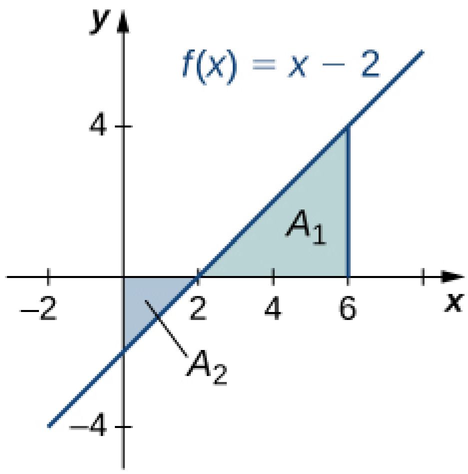
### Finding the Net Signed Area

Find the net signed area between the curve of the function $f(x) = 2x$ and the *x*-axis over the interval $[-3, 3]$.

[Show Solution]

**CHECKPOINT 1.9**

Find the net signed area of $f(x) = x - 2$ over the interval $[0, 6]$, illustrated in the following image.




## Total Area

One application of the definite integral is finding displacement when given a velocity function. If $v(t)$ represents the velocity of an object as a function of time, then the area under the curve tells us how far the object is from its original position. This is a very important application of the definite integral, and we examine it in more detail later in the chapter. For now, we're just going to look at some basics to get a feel for how this works by studying constant velocities.

When velocity is a constant, the area under the curve is just velocity times time. This idea is already very familiar. If a car travels away from its starting position in a straight line at a speed of 75 mph for 2 hours, then it is 150 mi away from its original position (Figure 1.20). Using integral notation, we have

$$\int_0^2 75 dt = 150.$$

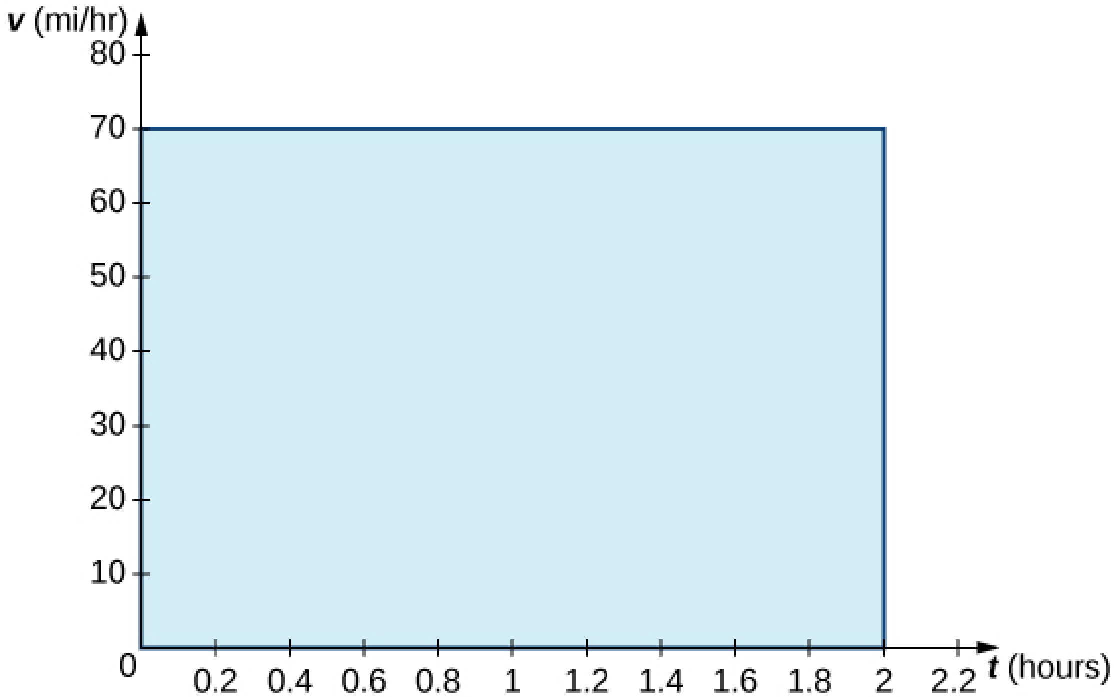



**Figure 1.20** The area under the curve $v(t) = 75$ tells us how far the car is from its starting point at a given time.

In the context of displacement, net signed area allows us to take direction into account. If a car travels straight north at a speed of 60 mph for 2 hours, it is 120 mi north of its starting position. If the car then turns around and travels south at a speed of 40 mph for 3 hours, it will be back at it starting position (Figure 1.21). Again, using integral notation, we have

$$\begin{aligned} \int_0^2 60dt + \int_2^5 -40dt &= 120 - 120 \\ &= 0. \end{aligned}$$

In this case the displacement is zero.

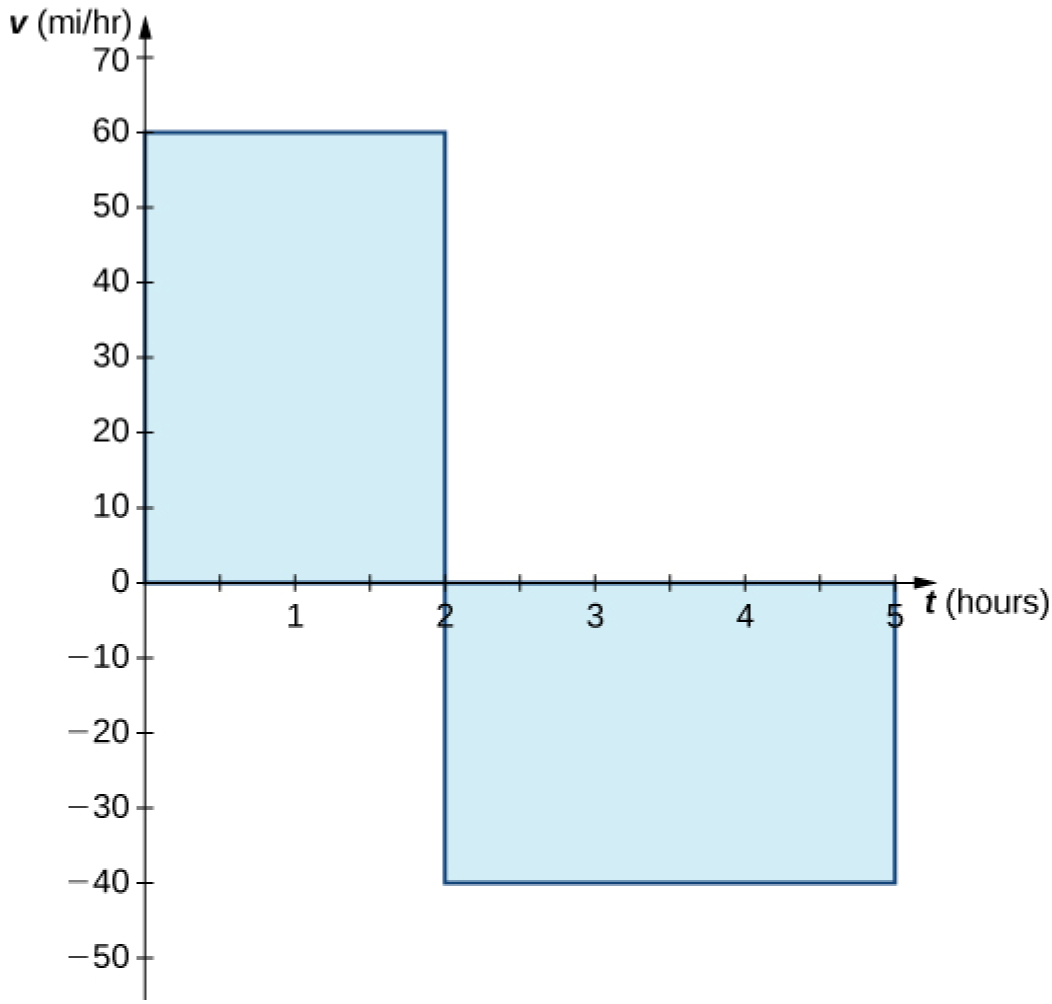



**Figure 1.21** The area above the axis and the area below the axis are equal, so the net signed area is zero.

Suppose we want to know how far the car travels overall, regardless of direction. In this case, we want to know the area between the curve and the *x*-axis, regardless of whether that area is above or below the axis. This is called the **total area**.

Graphically, it is easiest to think of calculating total area by adding the areas above the axis and the areas below the axis (rather than subtracting the areas below the axis, as we did with net signed area). To accomplish this mathematically, we use the absolute value function. Thus, the total distance traveled by the car is

$$\begin{aligned}\int_0^2 |60|\,dt + \int_2^5 |-40|\,dt &= \int_0^2 60\,dt + \int_2^5 40\,dt \\ &= 120 + 120 \\ &= 240.\end{aligned}$$

Bringing these ideas together formally, we state the following definitions.

#### DEFINITION

Let $f(x)$ be an integrable function defined on an interval $[a, b]$. Let $A_1$ represent the area between $f(x)$ and the $x$-axis that lies *above* the axis and let $A_2$ represent the area between $f(x)$ and the $x$-axis that lies *below* the axis. Then, the **net signed area** between $f(x)$ and the $x$-axis is given by

$$\int_a^b f(x)dx = A_1 - A_2.$$

The **total area** between $f(x)$ and the $x$-axis is given by

$$\int_a^b |f(x)|dx = A_1 + A_2.$$

#### EXAMPLE 1.10

### Finding the Total Area

Find the total area between $f(x) = x - 2$ and the $x$-axis over the interval $[0, 6]$.

[Show Solution]

#### CHECKPOINT 1.10

Find the total area between the function $f(x) = 2x$ and the $x$-axis over the interval $[-3, 3]$.

## Properties of the Definite Integral

The properties of indefinite integrals apply to definite integrals as well. Definite integrals also have properties that relate to the limits of integration. These properties, along with the rules of integration that we examine later in this chapter, help us manipulate expressions to evaluate definite integrals.

**RULE: PROPERTIES OF THE DEFINITE INTEGRAL**

1.

$$\int_a^a f(x)dx = 0 \tag{1.9}$$

If the limits of integration are the same, the integral is just a line and contains no area.

2.

$$\int_b^a f(x)dx = -\int_a^b f(x)dx \tag{1.10}$$

If the limits are reversed, then place a negative sign in front of the integral.

3.

$$\int_a^b [f(x) + g(x)]dx = \int_a^b f(x)dx + \int_a^b g(x)dx \tag{1.11}$$

The integral of a sum is the sum of the integrals.

4.

$$\int_a^b [f(x) - g(x)]dx = \int_a^b f(x)dx - \int_a^b g(x)dx \tag{1.12}$$

The integral of a difference is the difference of the integrals.

5.

$$\int_a^b cf(x)dx = c\int_a^b f(x) \tag{1.13}$$

for constant $c$. The integral of the product of a constant and a function is equal to the constant multiplied by the integral of the function.

6.

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx \tag{1.14}$$

Although this formula normally applies when *c* is between *a* and *b*, the formula holds for all values of *a*, *b*, and *c*, provided $f(x)$ is integrable on the largest interval.

**EXAMPLE 1.11**

### Using the Properties of the Definite Integral

Use the properties of the definite integral to express the definite integral of $f(x) = -3x^3 + 2x + 2$ over the interval $[-2, 1]$ as the sum of three definite integrals.

[Show Solution]

**CHECKPOINT 1.11**

Use the properties of the definite integral to express the definite integral of $f(x) = 6x^3 - 4x^2 + 2x - 3$ over the interval $[1, 3]$ as the sum of four definite integrals.

**EXAMPLE 1.12**

### Using the Properties of the Definite Integral

If it is known that $\int_0^8 f(x)dx = 10$ and $\int_0^5 f(x)dx = 5,$ find the value of $\int_5^8 f(x)dx.$

[Show Solution]

### CHECKPOINT 1.12

If it is known that $\int_1^5 f(x)dx = -3$ and $\int_2^5 f(x)dx = 4,$ find the value of $\int_1^2 f(x)dx.$

## Comparison Properties of Integrals

A picture can sometimes tell us more about a function than the results of computations. Comparing functions by their graphs as well as by their algebraic expressions can often give new insight into the process of integration. Intuitively, we might say that if a function $f(x)$ is above another function $g(x),$ then the area between $f(x)$ and the $x$-axis is greater than the area between $g(x)$ and the $x$-axis. This is true depending on the interval over which the comparison is made. The properties of definite integrals are valid whether $a < b, a = b,$ or $a > b.$ The following properties, however, concern only the case $a \le b,$ and are used when we want to compare the sizes of integrals.

### THEOREM 1.2

### Comparison Theorem

i. If $f(x) \ge 0$ for $a \le x \le b,$ then

$$\int_a^b f(x)dx \ge 0.$$

ii. If $f(x) \ge g(x)$ for $a \le x \le b,$ then

$$\int_a^b f(x)dx \ge \int_a^b g(x)dx.$$

iii. If *m* and *M* are constants such that $m \le f(x) \le M$ for $a \le x \le b,$ then

$$m(b - a) \le \int_a^b f(x)dx \le M(b - a).$$

### EXAMPLE 1.13

### Comparing Two Functions over a Given Interval

Compare $f(x) = \sqrt{1 + x^2}$ and $g(x) = \sqrt{1 + x}$ over the interval $[0, 1]$.

[Show Solution]

## Average Value of a Function

We often need to find the average of a set of numbers, such as an average test grade. Suppose you received the following test scores in your algebra class: 89, 90, 56, 78, 100, and 69. Your semester grade is your average of test scores and you want to know what grade to expect. We can find the average by adding all the scores and dividing by the number of scores. In this case, there are six test scores. Thus,

$$\frac{89 + 90 + 56 + 78 + 100 + 69}{6} = \frac{482}{6} \approx 80.33.$$

Therefore, your average test grade is approximately 80.33, which translates to a B– at most schools.

Suppose, however, that we have a function $v(t)$ that gives us the speed of an object at any time *t*, and we want to find the object's average speed. The function $v(t)$ takes on an infinite number of values, so we can't use the process just described. Fortunately, we can use a definite integral to find the average value of a function such as this.

Let $f(x)$ be continuous over the interval $[a, b]$ and let $[a, b]$ be divided into *n* subintervals of width $\Delta x = (b - a)/n.$ Choose a representative $x_i^*$ in each subinterval and calculate $f\left(x_i^*\right)$ for $i = 1, 2, \ldots, n.$ In other words, consider each $f\left(x_i^*\right)$ as a sampling of the function over each subinterval. The average value of the function may then be approximated as

$$\frac{f\left(x_1^*\right) + f\left(x_2^*\right) + \cdots + f\left(x_n^*\right)}{n},$$

which is basically the same expression used to calculate the average of discrete values.

But we know $\Delta x = \frac{b-a}{n},$ so $n = \frac{b-a}{\Delta x},$ and we get

$$\frac{f\left(x_1^*\right)+f\left(x_2^*\right)+\cdots+f\left(x_n^*\right)}{n}=\frac{f\left(x_1^*\right)+f\left(x_2^*\right)+\cdots+f\left(x_n^*\right)}{\frac{(b-a)}{\Delta x}}.$$

Following through with the algebra, the numerator is a sum that is represented as $\sum_{i=1}^{n} f\left(x_i^*\right)$, and we are dividing by a fraction. To divide by a fraction, invert the denominator and multiply. Thus, an approximate value for the average value of the function is given by

$$\begin{aligned}\frac{\sum_{i=1}^{n} f\left(x_i^*\right)}{\frac{(b-a)}{\Delta x}} &= \left(\frac{\Delta x}{b-a}\right)\sum_{i=1}^{n} f\left(x_i^*\right) \\ &= \left(\frac{1}{b-a}\right)\sum_{i=1}^{n} f\left(x_i^*\right)\Delta x.\end{aligned}$$

This is a Riemann sum. Then, to get the *exact* average value, take the limit as *n* goes to infinity. Thus, the average value of a function is given by

$$\frac{1}{b-a}\lim_{n\to\infty}\sum_{i=1}^{n} f\left(x_i\right)\Delta x=\frac{1}{b-a}\int_a^b f(x)dx.$$

**DEFINITION**

Let $f(x)$ be continuous over the interval $[a, b]$. Then, the **average value of the function** $f(x)$ (or $f_{\text{ave}}$) on $[a, b]$ is given by

$$f_{\text{ave}}=\frac{1}{b-a}\int_a^b f(x)dx.$$

**EXAMPLE 1.14**

### Finding the Average Value of a Linear Function

Find the average value of $f(x)=x+1$ over the interval $[0, 5]$.

[Show Solution]

CHECKPOINT 1.13

Find the average value of $f(x) = 6 - 2x$ over the interval $[0, 3]$.

## Section 1.2 Exercises

In the following exercises, express the limits as integrals.

60. $\lim_{n\to\infty} \sum_{i=1}^{n} \left(x_i^*\right)\Delta x$ over $[1, 3]$

61. $\lim_{n\to\infty} \sum_{i=1}^{n} \left(5\left(x_i^*\right)^2 - 3\left(x_i^*\right)^3\right)\Delta x$ over $[0, 2]$

62. $\lim_{n\to\infty} \sum_{i=1}^{n} \sin^2\left(2\pi x_i^*\right)\Delta x$ over $[0, 1]$

63. $\lim_{n\to\infty} \sum_{i=1}^{n} \cos^2\left(2\pi x_i^*\right)\Delta x$ over $[0, 1]$

In the following exercises, given $L_n$ or $R_n$ as indicated, express their limits as $n \to \infty$ as definite integrals, identifying the correct intervals.

64. $L_n = \frac{1}{n}\sum_{i=1}^{n} \frac{i-1}{n}$

65. $R_n = \frac{1}{n}\sum_{i=1}^{n} \frac{i}{n}$

66. $L_n = \frac{2}{n}\sum_{i=1}^{n} \left(1 + 2\frac{i-1}{n}\right)$

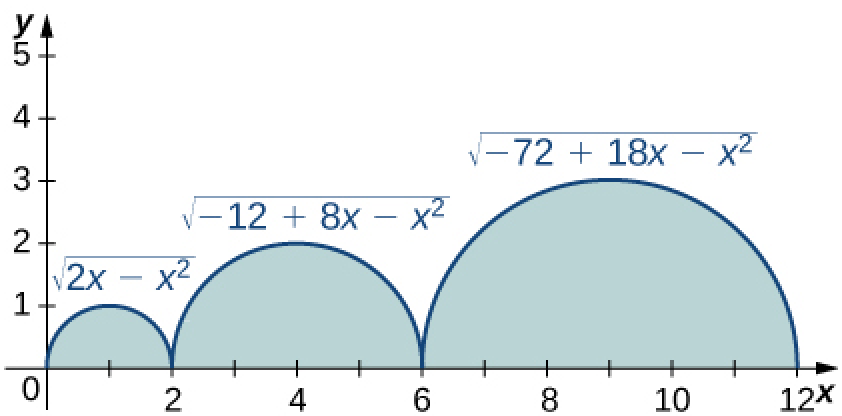
67. $R_n = \frac{3}{n}\sum_{i=1}^{n} \left(3 + 3\frac{i}{n}\right)$

68. $L_n = \frac{2\pi}{n}\sum_{i=1}^{n} 2\pi\frac{i-1}{n}\cos\left(2\pi\frac{i-1}{n}\right)$

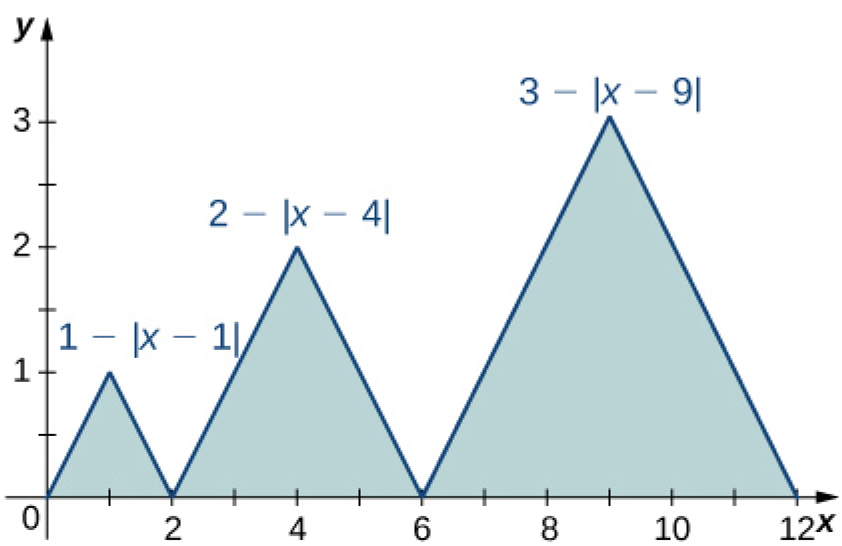
69. $R_n = \frac{1}{n}\sum_{i=1}^{n}\left(1+\frac{i}{n}\right)\log\left(\left(1+\frac{i}{n}\right)^2\right)$

In the following exercises, evaluate the integrals of the functions graphed using the formulas for areas of triangles and circles, and subtracting the areas below the *x*-axis.

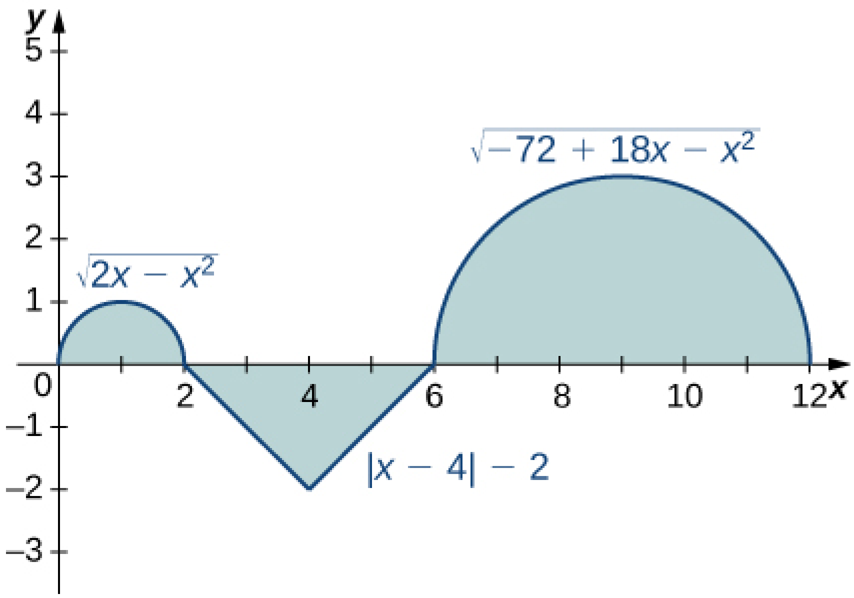
70.



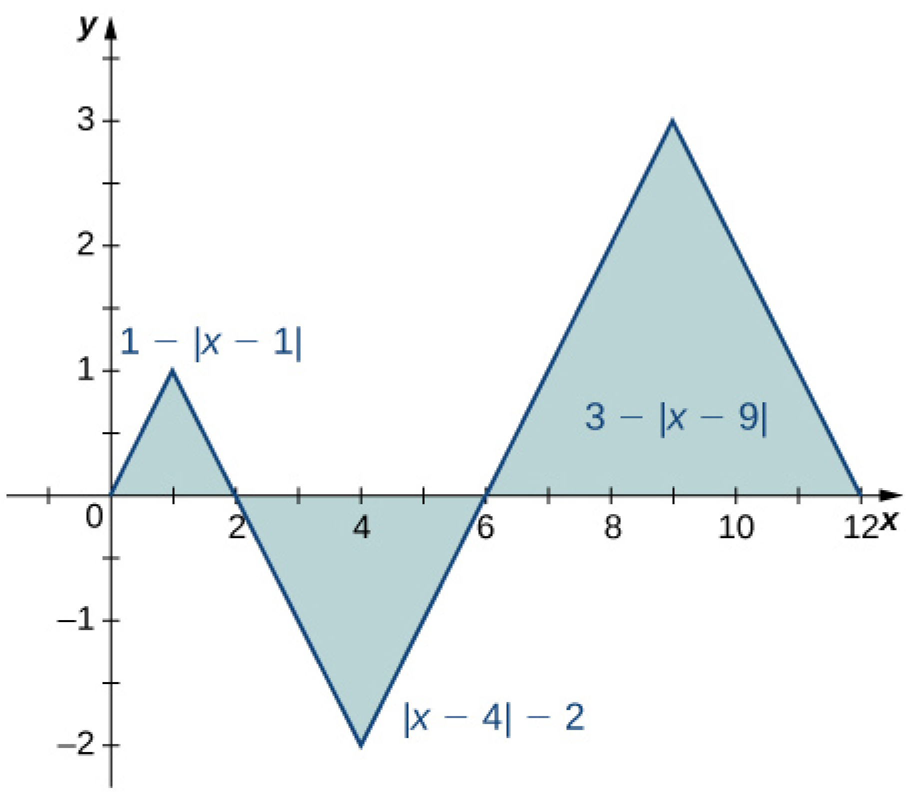

71.



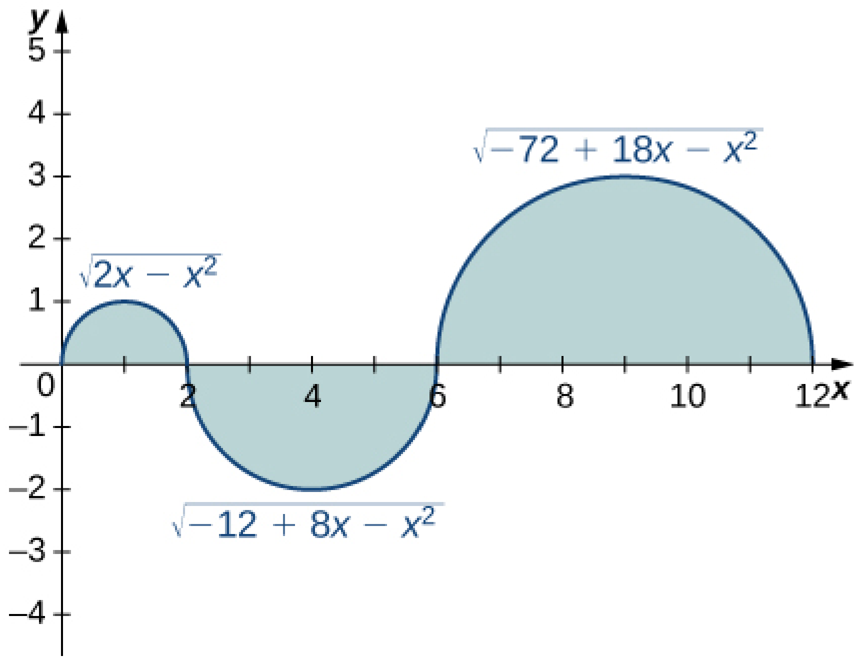

72.



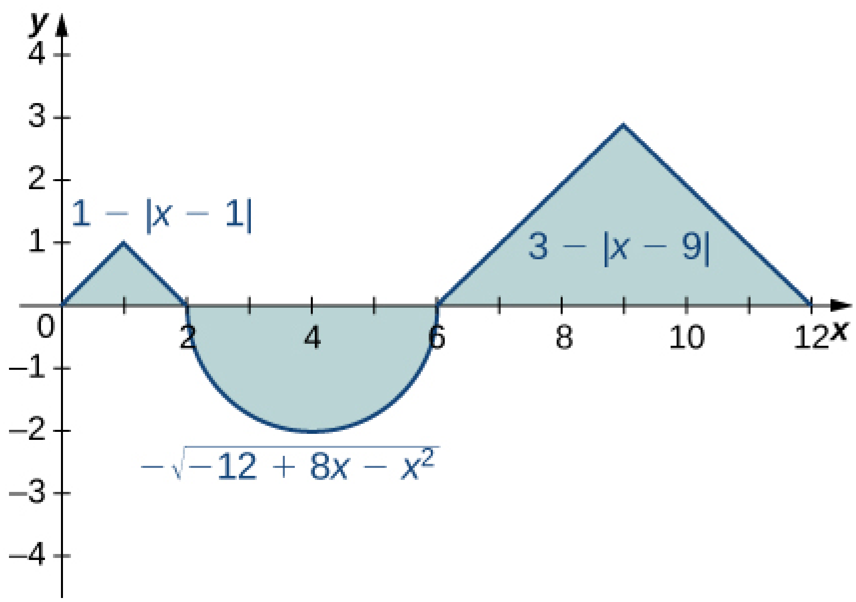

73.




74.




75.




In the following exercises, evaluate the integral using area formulas.

76. $\int_0^3 (3 - x)dx$

77. $\int_2^3 (3 - x)dx$

78. $\int_{-3}^{3}(3-|x|)dx$

79. $\int_{0}^{6}(3-|x-3|)dx$

80. $\int_{-2}^{2}\sqrt{4-x^2}dx$

81. $\int_{1}^{5}\sqrt{4-(x-3)^2}dx$

82. $\int_{0}^{12}\sqrt{36-(x-6)^2}dx$

83. $\int_{-2}^{3}(3-|x|)dx$

In the following exercises, use averages of values at the left (*L*) and right (*R*) endpoints to compute the integrals of the piecewise linear functions with graphs that pass through the given list of points over the indicated intervals.

84. $\{(0, 0), (2, 1), (4, 3), (5, 0), (6, 0), (8, 3)\}$ over $[0, 8]$

85. $\{(0, 2), (1, 0), (3, 5), (5, 5), (6, 2), (8, 0)\}$ over $[0, 8]$

86. $\{(-4, -4), (-2, 0), (0, -2), (3, 3), (4, 3)\}$ over $[-4, 4]$

87. $\{(-4, 0), (-2, 2), (0, 0), (1, 2), (3, 2), (4, 0)\}$ over $[-4, 4]$

Suppose that $\int_{0}^{4}f(x)dx = 5$ and $\int_{0}^{2}f(x)dx = -3,$ and $\int_{0}^{4}g(x)dx = -1$ and $\int_{0}^{2}g(x)dx = 2.$ In the following exercises, compute the integrals.

88. $\int_{0}^{4}(f(x)+g(x))dx$

89. $\int_{2}^{4}(f(x)+g(x))dx$

90. $\int_{0}^{2}(f(x)-g(x))dx$

91. $\int_{2}^{4}(f(x)-g(x))dx$

92. $\int_{0}^{2}(3f(x)-4g(x))dx$

93. $\int_{2}^{4}(4f(x)-3g(x))dx$

In the following exercises, use the identity $\int_{-A}^{A}f(x)dx = \int_{-A}^{0}f(x)dx + \int_{0}^{A}f(x)dx$ to compute the integrals.

94. $\int_{-\pi}^{\pi} \frac{\sin t}{1+t^2} dt$ $(Hint\text{: } \sin(-t) = -\sin(t))$

95. $\int_{-\sqrt{\pi}}^{\sqrt{\pi}} \frac{t}{1+\cos t} dt$

In the following exercises, find the net signed area between $f(x)$ and the x-axis.

96. $\int_1^3 (2-x)dx$ (*Hint:* Look at the graph of *f*.)

97. $\int_2^4 (x-3)^3 dx$ (*Hint:* Look at the graph of *f*.)

In the following exercises, given that $\int_0^1 x dx = \frac{1}{2}, \int_0^1 x^2 dx = \frac{1}{3},$ and $\int_0^1 x^3 dx = \frac{1}{4},$ compute the integrals.

98. $\int_0^1 \left(1 + x + x^2 + x^3\right) dx$

99. $\int_0^1 \left(1 - x + x^2 - x^3\right) dx$

100. $\int_0^1 (1-x)^2 dx$

101. $\int_0^1 (1-2x)^3 dx$

102. $\int_0^1 \left(6x - \frac{4}{3}x^2\right) dx$

103. $\int_0^1 \left(7 - 5x^3\right) dx$

In the following exercises, use the comparison theorem.

104. Show that $\int_0^3 \left(x^2 - 6x + 9\right) dx \geq 0.$

105. Show that $\int_{-2}^3 (x-3)(x+2)dx \leq 0.$

106. Show that $\int_0^1 \sqrt{1+x^3} dx \leq \int_0^1 \sqrt{1+x^2} dx.$

107. Show that $\int_1^2 \sqrt{1+x} dx \leq \int_1^2 \sqrt{1+x^2} dx.$

108. Show that $\int_0^{\pi/2} \sin t dt \geq \frac{\pi}{4}.$ $(Hint\text{: } \sin t \geq \frac{2t}{\pi}$ over $\left[0, \frac{\pi}{2}\right])$

109. Show that $\int_{-\pi/4}^{\pi/4} \cos t\,dt \geq \pi\sqrt{2}/4$.

In the following exercises, find the average value $f_{\text{ave}}$ of $f$ between $a$ and $b$, and find a point $c$, where $f(c) = f_{\text{ave}}$.

110. $f(x) = x^2,\ a = -1,\ b = 1$

111. $f(x) = x^5,\ a = -1,\ b = 1$

112. $f(x) = \sqrt{4 - x^2},\ a = 0,\ b = 2$

113. $f(x) = (3 - |x|),\ a = -3,\ b = 3$

114. $f(x) = \sin x,\ a = 0,\ b = 2\pi$

115. $f(x) = \cos x,\ a = 0,\ b = 2\pi$

In the following exercises, approximate the average value using Riemann sums $L_{100}$ and $R_{100}$. How does your answer compare with the exact given answer?

116. **[T]** $y = \ln(x)$ over the interval $[1, 4]$; the exact solution is $\frac{\ln(256)}{3} - 1$.

117. **[T]** $y = e^{x/2}$ over the interval $[0, 1]$; the exact solution is $2\left(\sqrt{e} - 1\right)$.

118. **[T]** $y = \tan x$ over the interval $\left[0, \frac{\pi}{4}\right]$; the exact solution is $\frac{2\ln(2)}{\pi}$.

119. **[T]** $y = \frac{x+1}{\sqrt{4-x^2}}$ over the interval $[-1, 1]$; the exact solution is $\frac{\pi}{6}$.

In the following exercises, compute the average value using the left Riemann sums $L_N$ for $N = 1, 10, 100$. How does the accuracy compare with the given exact value?

120. **[T]** $y = x^2 - 4$ over the interval $[0, 2]$; the exact solution is $-\frac{8}{3}$.

121. **[T]** $y = xe^{x^2}$ over the interval $[0, 2]$; the exact solution is $\frac{1}{4}\left(e^4 - 1\right)$.

122. **[T]** $y = \left(\frac{1}{2}\right)^x$ over the interval $[0, 4]$; the exact solution is $\frac{15}{64\ln(2)}$.

123. **[T]** $y = x\sin\left(x^2\right)$ over the interval $[-\pi, 0]$; the exact solution is $\frac{\cos\left(\pi^2\right) - 1}{2\pi}$.

124. Suppose that $A = \int_0^{2\pi} \sin^2 t dt$ and $B = \int_0^{2\pi} \cos^2 t dt$. Show that $A + B = 2\pi$ and $A = B$.

125. Suppose that $A = \int_{-\pi/4}^{\pi/4} \sec^2 t dt = \pi$ and $B = \int_{-\pi/4}^{\pi/4} \tan^2 t dt$. Show that $A - B = \frac{\pi}{2}$.

126. Show that the average value of $\sin^2 t$ over $[0, 2\pi]$ is equal to 1/2 Without further calculation, determine whether the average value of $\sin^2 t$ over $[0, \pi]$ is also equal to 1/2.

127. Show that the average value of $\cos^2 t$ over $[0, 2\pi]$ is equal to $1/2$. Without further calculation, determine whether the average value of $\cos^2(t)$ over $[0, \pi]$ is also equal to $1/2$.

128. Explain why the graphs of a quadratic function (parabola) $p(x)$ and a linear function $\ell(x)$ can intersect in at most two points. Suppose that $p(a) = \ell(a)$ and $p(b) = \ell(b)$, and that $\int_a^b p(t)dt > \int_a^b \ell(t)dt$. Explain why $\int_c^d p(t) > \int_c^d \ell(t)dt$ whenever $a \le c < d \le b$.

129. Suppose that parabola $p(x) = ax^2 + bx + c$ opens downward $(a < 0)$ and has a vertex of $y = \frac{-b}{2a} > 0$. For which interval $[A, B]$ is $\int_A^B \left(ax^2 + bx + c\right)dx$ as large as possible?

130. Suppose $[a, b]$ can be subdivided into subintervals $a = a_0 < a_1 < a_2 < \cdots < a_N = b$ such that either $f \ge 0$ over $\left[a_{i-1}, a_i\right]$ or $f \le 0$ over $\left[a_{i-1}, a_i\right]$. Set $A_i = \int_{a_{i-1}}^{a_i} f(t)dt$.

a. Explain why $\int_a^b f(t)dt = A_1 + A_2 + \cdots + A_N$.

b. Then, explain why $\left|\int_a^b f(t)dt\right| \le \int_a^b |f(t)|dt$.

131. Suppose *f* and *g* are continuous functions such that $\int_c^d f(t)dt \le \int_c^d g(t)dt$ for every subinterval $[c, d]$ of $[a, b]$. Explain why $f(x) \le g(x)$ for all values of *x*.

132. Suppose the average value of *f* over $[a, b]$ is 1 and the average value of *f* over $[b, c]$ is 1 where $a < c < b$. Show that the average value of *f* over $[a, c]$ is also 1.

133. Suppose that $[a, b]$ can be partitioned. taking $a = a_0 < a_1 < \cdots < a_N = b$ such that the average value of *f* over each subinterval $\left[a_{i-1}, a_i\right] = 1$ is equal to 1 for each $i = 1, \ldots, N$. Explain why the average value of *f* over $[a, b]$ is also equal to 1.

134. Suppose that for each *i* such that $1 \le i \le N$ one has $\int_{i-1}^{i} f(t)dt = i$. Show that $\int_0^N f(t)dt = \frac{N(N+1)}{2}$.

135. Suppose that for each *i* such that $1 \le i \le N$ one has $\int_{i-1}^{i} f(t)dt = i^2$. Show that $\int_0^N f(t)dt = \frac{N(N+1)(2N+1)}{6}$.

136. **[T]** Compute the left and right Riemann sums $L_{10}$ and $R_{10}$ and their average $\frac{L_{10}+R_{10}}{2}$ for $f(t) = t^2$ over $[0, 1]$. Given that $\int_0^1 t^2 dt = 0.3\overline{3}$, to how many decimal places is $\frac{L_{10}+R_{10}}{2}$ accurate?

137. **[T]** Compute the left and right Riemann sums, $L_{10}$ and $R_{10}$, and their average $\frac{L_{10}+R_{10}}{2}$ for $f(t) = \left(4 - t^2\right)$ over $[1, 2]$. Given that $\int_1^2 \left(4 - t^2\right)dt = 1.6\overline{6}$, to how many decimal places is $\frac{L_{10}+R_{10}}{2}$ accurate?

138. If $\int_1^5 \sqrt{1 + t^4}dt = 41.7133...$, what is $\int_1^5 \sqrt{1 + u^4}du$?

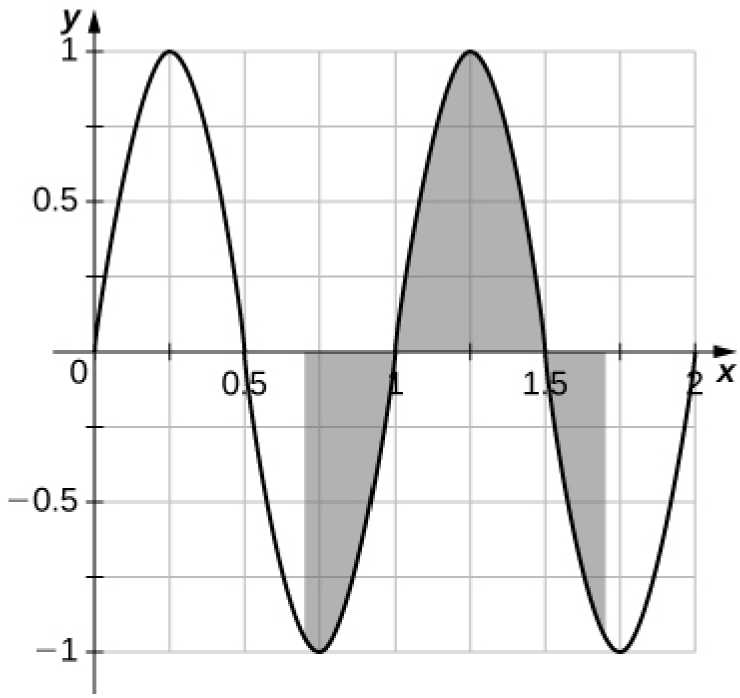
139. Estimate $\int_0^1 tdt$ using the left and right endpoint sums, each with a single rectangle. How does the average of these left and right endpoint sums compare with the actual value $\int_0^1 tdt$?

140. Estimate $\int_0^1 tdt$ by comparison with the area of a single rectangle with height equal to the value of *t* at the midpoint $t = \frac{1}{2}$. How does this midpoint estimate compare with the actual value $\int_0^1 tdt$?

141. From the graph of $\sin(2\pi x)$ shown:

a. Explain why $\int_0^1 \sin(2\pi t)dt = 0$.

b. Explain why, in general, $\int_a^{a+1} \sin(2\pi t)dt = 0$ for any value of *a*.




142. If *f* is 1-periodic $(f(t+1)=f(t))$, odd, and integrable over $[0, 1]$, is it always true that $\int_0^1 f(t)dt = 0$?

143. If *f* is 1-periodic and $\int_0^1 f(t)dt = A$, is it necessarily true that $\int_a^{1+a} f(t)dt = A$ for all *A*?

## Learning Objectives

1.3.1. Describe the meaning of the Mean Value Theorem for Integrals.
1.3.2. State the meaning of the Fundamental Theorem of Calculus, Part 1.
1.3.3. Use the Fundamental Theorem of Calculus, Part 1, to evaluate derivatives of integrals.
1.3.4. State the meaning of the Fundamental Theorem of Calculus, Part 2.
1.3.5. Use the Fundamental Theorem of Calculus, Part 2, to evaluate definite integrals.
1.3.6. Explain the relationship between differentiation and integration.

In the previous two sections, we looked at the definite integral and its relationship to the area under the curve of a function. Unfortunately, so far, the only tools we have available to calculate the value of a definite integral are geometric area formulas and limits of Riemann sums, and both approaches are extremely cumbersome. In this section we look at some more powerful and useful techniques for evaluating definite integrals.

These new techniques rely on the relationship between differentiation and integration. This relationship was discovered and explored by both Sir Isaac Newton and Gottfried Wilhelm Leibniz (among others) during the late 1600s and early 1700s, and it is codified in what we now call the **Fundamental Theorem of Calculus**, which has two parts that we examine in this section. Its very name indicates how central this theorem is to the entire development of calculus.

**MEDIA**

Isaac Newton's contributions to mathematics and physics changed the way we look at the world. The relationships he discovered, codified as Newton's laws and the law of universal gravitation, are still taught as foundational material in physics today, and his calculus has spawned entire fields of mathematics. To learn more, read a brief biography of Newton with multimedia clips.

Before we get to this crucial theorem, however, let's examine another important theorem, the Mean Value Theorem for Integrals, which is needed to prove the Fundamental Theorem of Calculus.

## The Mean Value Theorem for Integrals

The **Mean Value Theorem for Integrals** states that a continuous function on a closed interval takes on its average value at some point in that interval. The theorem guarantees that if $f(x)$ is continuous, a point *c* exists in an interval $[a, b]$ such that the value of the function at *c* is equal to the average value of $f(x)$ over $[a, b]$. We state this theorem mathematically with the help of the formula for the average value of a function that we presented at the end of the preceding section.

**THEOREM 1.3**

### The Mean Value Theorem for Integrals

If $f(x)$ is continuous over an interval $[a, b]$ , then there is at least one point $c \in [a, b]$ such that

$$f(c) = \frac{1}{b-a} \int_a^b f(x)\, dx. \tag{1.15}$$

This formula can also be stated as

$$\int_a^b f(x)\, dx = f(c)\,(b-a).$$

### Proof

Since $f(x)$ is continuous on $[a, b]$ , by the extreme value theorem (see Maxima and Minima), it assumes minimum and maximum values—*m* and *M*, respectively—on $[a, b]$ . Then, for all *x* in $[a, b]$ , we have $m \le f(x) \le M$. Therefore, by the comparison theorem (see The Definite Integral), we have

$$m\,(b-a) \le \int_a^b f(x)\, dx \le M\,(b-a).$$

Dividing by $b - a$ gives us

$$m \le \frac{1}{b-a} \int_a^b f(x)\, dx \le M.$$

Since $\frac{1}{b-a} \int_a^b f(x)\, dx$ is a number between *m* and *M*, and since $f(x)$ is continuous and assumes the values *m* and *M* over $[a, b]$ , by the Intermediate Value Theorem (see Continuity), there is a number *c* over $[a, b]$ such that

$$f(c) = \frac{1}{b-a} \int_a^b f(x)\, dx,$$

and the proof is complete.

□

**EXAMPLE 1.15**

### Finding the Average Value of a Function

Find the average value of the function $f(x) = 8 - 2x$ over the interval $[0, 4]$ and find *c* such that $f(c)$ equals the average value of the function over $[0, 4]$.

[Show Solution]

**CHECKPOINT 1.14**

Find the average value of the function $f(x) = \frac{x}{2}$ over the interval $[0, 6]$ and find *c* such that $f(c)$ equals the average value of the function over $[0, 6]$.

**EXAMPLE 1.16**

### Finding the Point Where a Function Takes on Its Average Value

Given $\int_0^3 x^2 dx = 9$, find *c* such that $f(c)$ equals the average value of $f(x) = x^2$ over $[0, 3]$.

[Show Solution]

**CHECKPOINT 1.15**

Given $\int_0^3 \left(2x^2 - 1\right) dx = 15$, find *c* such that $f(c)$ equals the average value of $f(x) = 2x^2 - 1$ over $[0, 3]$.

## Fundamental Theorem of Calculus Part 1: Integrals and Antiderivatives

As mentioned earlier, the Fundamental Theorem of Calculus is an extremely powerful theorem that establishes the relationship between differentiation and integration, and gives us a way to evaluate definite integrals without using Riemann sums or calculating areas. The theorem is comprised of two parts, the first of which, the **Fundamental Theorem of Calculus, Part 1**, is stated here. Part 1 establishes the relationship between differentiation and integration.

**THEOREM 1.4**

### Fundamental Theorem of Calculus, Part 1

If $f(x)$ is continuous over an interval $[a, b]$, and the function $F(x)$ is defined by

$$F(x) = \int_a^x f(t)\, dt, \tag{1.16}$$

then $F'(x) = f(x)$ over $[a, b]$.

Before we delve into the proof, a couple of subtleties are worth mentioning here. First, a comment on the notation. Note that we have defined a function, $F(x)$, as the definite integral of another function, $f(t)$, from the point *a* to the point *x*. At first glance, this is confusing, because we have said several times that a definite integral is a number, and here it looks like it's a function. The key here is to notice that for any particular value of *x*, the definite integral is a number. So the function $F(x)$ returns a number (the value of the definite integral) for each value of *x*.

Second, it is worth commenting on some of the key implications of this theorem. There is a reason it is called the *Fundamental* Theorem of Calculus. Not only does it establish a relationship between integration and differentiation, but also it guarantees that any integrable function has an antiderivative. Specifically, it guarantees that any continuous function has an antiderivative.

## Proof

Applying the definition of the derivative, we have

$$\begin{aligned} F'(x) &= \lim_{h \to 0} \frac{F(x+h)-F(x)}{h} \\ &= \lim_{h \to 0} \frac{1}{h}\left[\int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt\right] \\ &= \lim_{h \to 0} \frac{1}{h}\left[\int_a^{x+h} f(t)\,dt + \int_x^a f(t)\,dt\right] \\ &= \lim_{h \to 0} \frac{1}{h}\int_x^{x+h} f(t)\,dt. \end{aligned}$$

Looking carefully at this last expression, we see $\frac{1}{h}\int_x^{x+h} f(t)\,dt$ is just the average value of the function $f(x)$ over the interval $[x, x+h]$. Therefore, by [The Mean Value Theorem for Integrals](), there is some number *c* in $[x, x+h]$ such that

$$\frac{1}{h}\int_x^{x+h} f(x)\,dx = f(c).$$

In addition, since *c* is between *x* and *x* + *h*, *c* approaches *x* as *h* approaches zero. Also, since $f(x)$ is continuous, we have $\lim_{h \to 0} f(c) = \lim_{c \to x} f(c) = f(x)$. Putting all these pieces together, we have

$$\begin{aligned} F'(x) &= \lim_{h \to 0} \frac{1}{h} \int_x^{x+h} f(x)\, dx \\ &= \lim_{h \to 0} f(c) \\ &= f(x), \end{aligned}$$

and the proof is complete.

□

## EXAMPLE 1.17

### Finding a Derivative with the Fundamental Theorem of Calculus

Use the Fundamental Theorem of Calculus, Part 1 to find the derivative of

$$g(x) = \int_1^x \frac{1}{t^3 + 1}\, dt.$$

[Show Solution]

## CHECKPOINT 1.16

Use the Fundamental Theorem of Calculus, Part 1 to find the derivative of $g(r) = \int_0^r \sqrt{x^2 + 4}\, dx.$

## EXAMPLE 1.18

### Using the Fundamental Theorem and the Chain Rule to Calculate Derivatives

Let $F(x) = \int_1^{\sqrt{x}} \sin t\,dt$. Find $F'(x)$.

[Show Solution]

**CHECKPOINT 1.17**

Let $F(x) = \int_1^{x^3} \cos t\,dt$. Find $F'(x)$.

**EXAMPLE 1.19**

### Using the Fundamental Theorem of Calculus with Two Variable Limits of Integration

Let $F(x) = \int_x^{2x} t^3\,dt$. Find $F'(x)$.

[Show Solution]

**CHECKPOINT 1.18**

Let $F(x) = \int_x^{x^2} \cos t\,dt$. Find $F'(x)$.

## Fundamental Theorem of Calculus, Part 2: The Evaluation Theorem

The Fundamental Theorem of Calculus, Part 2, is perhaps the most important theorem in calculus. After tireless efforts by mathematicians for approximately 500 years, new techniques emerged that provided scientists with the necessary tools to explain many phenomena. Using calculus, astronomers could finally determine distances in space and map planetary orbits. Everyday financial problems such as calculating marginal costs or predicting total profit could now be handled with simplicity and accuracy. Engineers could calculate the bending strength of materials or the three-dimensional motion of objects. Our view of the world was forever changed with calculus.

After finding approximate areas by adding the areas of *n* rectangles, the application of this theorem is straightforward by comparison. It almost seems too simple that the area of an entire curved region can be calculated by just evaluating an antiderivative at the first and last endpoints of an interval.

**THEOREM 1.5**

**The Fundamental Theorem of Calculus, Part 2**

If *f* is continuous over the interval $[a, b]$ and $F(x)$ is any antiderivative of $f(x)$, then

$$\int_a^b f(x)\,dx = F(b) - F(a). \tag{1.17}$$

We often see the notation $F(x)|_a^b$ to denote the expression $F(b) - F(a)$. We use this vertical bar and associated limits *a* and *b* to indicate that we should evaluate the function $F(x)$ at the upper limit (in this case, *b*), and subtract the value of the function $F(x)$ evaluated at the lower limit (in this case, *a*).

The **Fundamental Theorem of Calculus, Part 2** (also known as the **evaluation theorem**) states that if we can find an antiderivative for the integrand, then we can evaluate the definite integral by evaluating the antiderivative at the endpoints of the interval and subtracting.

### Proof

Let $P = \{x_i\}, i = 0, 1, \ldots, n$ be a regular partition of $[a, b]$. Then, we can write

$$\begin{aligned} F(b) - F(a) &= F(x_n) - F(x_0) \\ &= \left[F(x_n) - F\left(x_{n-1}\right)\right] + \left[F\left(x_{n-1}\right) - F\left(x_{n-2}\right)\right] + \ldots + \left[F(x_1) - \right. \\ &= \sum_{i=1}^{n} \left[F(x_i) - F\left(x_{i-1}\right)\right]. \end{aligned}$$

Now, we know *F* is an antiderivative of *f* over $[a, b]$, so by the Mean Value Theorem (see The Mean Value Theorem) for $i = 0, 1, \ldots, n$ we can find $c_i$ in $\left[x_{i-1}, x_i\right]$ such that

$$F(x_i) - F\left(x_{i-1}\right) = F'(c_i)\left(x_i - x_{i-1}\right) = f(c_i)\,\Delta x.$$

Then, substituting into the previous equation, we have

$$F(b) - F(a) = \sum_{i=1}^{n} f(c_i)\,\Delta x.$$

Taking the limit of both sides as $n \to \infty$, we obtain

$$\begin{aligned} F(b) - F(a) &= \lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\,\Delta x \\ &= \int_a^b f(x)\,dx. \end{aligned}$$

□

**EXAMPLE 1.20**

### Evaluating an Integral with the Fundamental Theorem of Calculus

Use The Fundamental Theorem of Calculus, Part 2 to evaluate

$$\int_{-2}^{2} \left(t^2 - 4\right) dt.$$

[Show Solution]

EXAMPLE 1.21

### Evaluating a Definite Integral Using the Fundamental Theorem of Calculus, Part 2

Evaluate the following integral using the Fundamental Theorem of Calculus, Part 2:

$$\int_1^9 \frac{x-1}{\sqrt{x}}\,dx.$$

[Show Solution]

CHECKPOINT 1.19

Use The Fundamental Theorem of Calculus, Part 2 to evaluate $\int_1^2 x^{-4}\,dx.$

EXAMPLE 1.22

### A Roller-Skating Race

James and Kathy are racing on roller skates. They race along a long, straight track, and whoever has gone the farthest after 5 sec wins a prize. If James can skate at a velocity of $f(t) = 5 + 2t$ ft/sec and Kathy can skate at a velocity of $g(t) = 10 + \cos\left(\frac{\pi}{2}t\right)$ ft/sec, who is going to win the race?

[Show Solution]

CHECKPOINT 1.20

Suppose James and Kathy have a rematch, but this time the official stops the contest after only 3 sec. Does this change the outcome?

STUDENT PROJECT

## A Parachutist in Free Fall

**Figure 1.30** Skydivers can adjust the velocity of their dive by changing the position of their body during the free fall. (credit: Jeremy T. Lock)

Julie is an avid skydiver. She has more than 300 jumps under her belt and has mastered the art of making adjustments to her body position in the air to control how fast she falls. If she arches her back and points her belly toward the ground, she reaches a terminal velocity of approximately 120 mph (176 ft/sec). If, instead, she orients her body with her head straight down, she falls faster, reaching a terminal velocity of 150 mph (220 ft/sec).

Since Julie will be moving (falling) in a downward direction, we assume the downward direction is positive to simplify our calculations. Julie executes her jumps from an altitude of 12,500 ft. After she exits the aircraft, she immediately starts falling at a velocity given by $v(t) = 32t$. She continues to accelerate according to this velocity function until she reaches terminal velocity. After she reaches terminal velocity, her speed remains constant until she pulls her ripcord and slows down to land.

On her first jump of the day, Julie orients herself in the slower "belly down" position (terminal velocity is 176 ft/sec). Using this information, answer the following questions.

1. How long after she exits the aircraft does Julie reach terminal velocity?
2. Based on your answer to question 1, set up an expression involving one or more integrals that represents the distance Julie falls after 30 sec.
3. If Julie pulls her ripcord at an altitude of 3000 ft, how long does she spend in a free fall?
4. Julie pulls her ripcord at 3000 ft. It takes 5 sec for her parachute to open completely and for her to slow down, during which time she falls another 400 ft. After her canopy is fully open, her speed is reduced to 16 ft/sec. Find the total time Julie spends in the air, from the time she leaves the airplane until the time her feet touch the ground.
   On Julie's second jump of the day, she decides she wants to fall a little faster and orients herself in the "head down" position. Her terminal velocity in this position is 220 ft/sec. Answer these questions based on this velocity:
5. How long does it take Julie to reach terminal velocity in this case?
6. Before pulling her ripcord, Julie reorients her body in the "belly down" position so she is not moving quite as fast when her parachute opens. If she begins this maneuver at an altitude of 4000 ft, how long does she spend in a free fall before beginning the reorientation?
   Some jumpers wear "wingsuits" (see Figure 1.31). These suits have fabric panels between the arms and legs and allow the wearer to glide around in a free fall, much like a flying squirrel. (Indeed, the suits are sometimes called "flying squirrel suits.") When wearing these suits, terminal velocity can be reduced to about 30 mph (44 ft/sec), allowing the wearers a much longer time in the air. Wingsuit flyers still use parachutes to land; although the vertical velocities are within the margin of safety, horizontal velocities can exceed 70 mph, much too fast to land safely.

**Figure 1.31** The fabric panels on the arms and legs of a wingsuit work to reduce the vertical velocity of a skydiver's fall. (credit: Richard Schneider)

Answer the following question based on the velocity in a wingsuit.

7. If Julie dons a wingsuit before her third jump of the day, and she pulls her ripcord at an altitude of 3000 ft, how long does she get to spend gliding around in the air?

## Section 1.3 Exercises

144. Consider two athletes running at variable speeds $v_1(t)$ and $v_2(t)$. The runners start and finish a race at exactly the same time. Explain why the two runners must be going the same speed at some point.

145. Two mountain climbers start their climb at base camp, taking two different routes, one steeper than the other, and arrive at the peak at exactly the same time. Is it necessarily true that, at some point, both climbers increased in altitude at the same rate?

146. To get on a certain toll road a driver has to take a card that lists the mile entrance point. The card also has a timestamp. When going to pay the toll at the exit, the driver is surprised to receive a speeding ticket along with the toll. Explain how this can happen.

147. Set $F(x) = \int_1^x (1-t)\,dt$. Find $F'(2)$ and the average value of $F'$ over $[1,2]$.

In the following exercises, use the Fundamental Theorem of Calculus, Part 1, to find each derivative.

148. $\frac{d}{dx}\int_1^x e^{-t^2}\,dt$

149. $\frac{d}{dx}\int_1^x e^{\cos t}\,dt$

150. $\frac{d}{dx}\int_3^x \sqrt{9-y^2}\,dy$

151. $\frac{d}{dx}\int_4^x \frac{ds}{\sqrt{16-s^2}}$

152. $\frac{d}{dx}\int_x^{2x} t\,dt$

153. $\frac{d}{dx}\int_0^{\sqrt{x}} t\,dt$

154. $\frac{d}{dx}\int_0^{\sin x} \sqrt{1-t^2}\,dt$

155. $\frac{d}{dx}\int_{\cos x}^{1} \sqrt{1-t^2}\,dt$

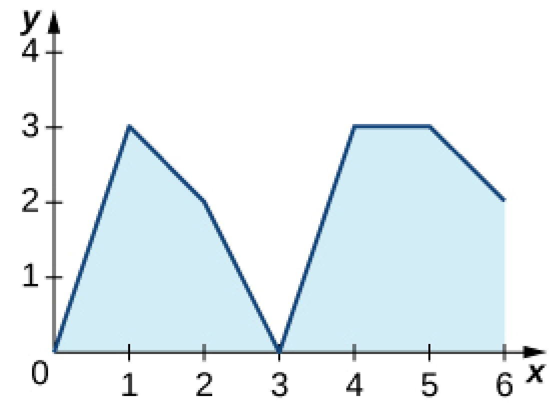
156. $\frac{d}{dx}\int_1^{\sqrt{x}} \frac{t^2}{1+t^4}\,dt$

157. $\frac{d}{dx}\int_1^{x^2} \frac{\sqrt{t}}{1+t}\,dt$

158. $\frac{d}{dx}\int_0^{\ln x} e^t\,dt$

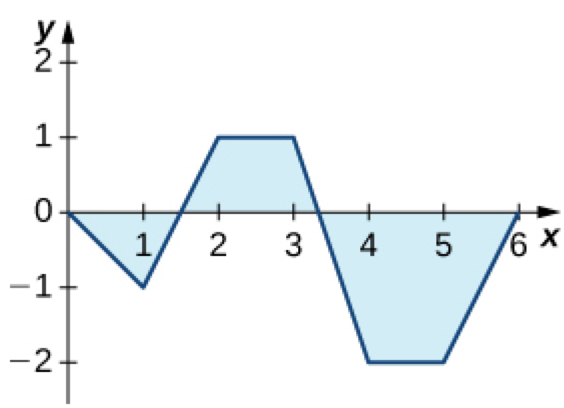
159. $\frac{d}{dx}\int_1^{e^x} \ln u^2\,du$

160. The graph of $y = \int_0^x f(t)dt$, where *f* is a piecewise constant function, is shown here.



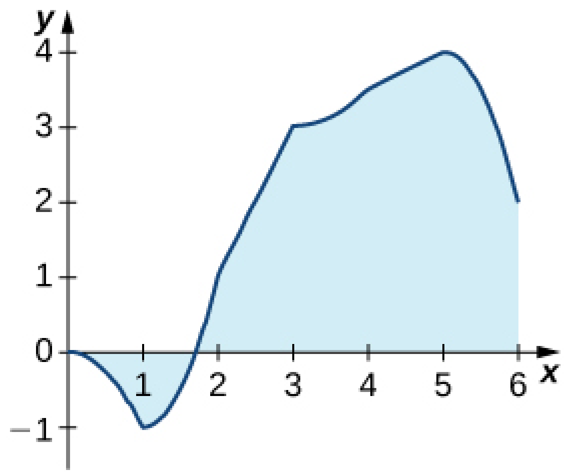

a. Over which intervals is *f* positive? Over which intervals is it negative? Over which intervals, if any, is it equal to zero?
b. What are the maximum and minimum values of *f*?
c. What is the average value of *f*?

161. The graph of $y = \int_0^x f(t)dt$, where *f* is a piecewise constant function, is shown here.



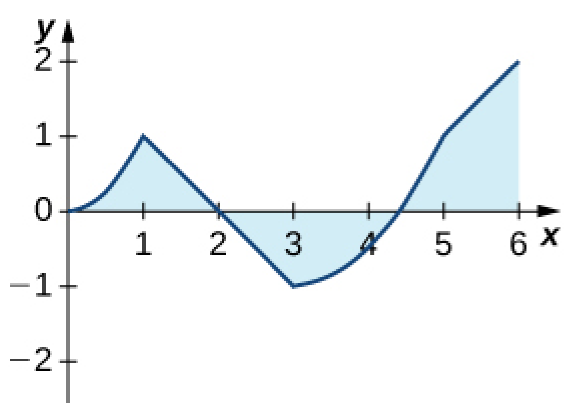

a. Over which intervals is *f* positive? Over which intervals is it negative? Over which intervals, if any, is it equal to zero?
b. What are the maximum and minimum values of *f*?
c. What is the average value of *f*?

162. The graph of $y = \int_0^x \ell(t)dt,$ where $\ell$ is a piecewise linear function, is shown here.




a. Over which intervals is $\ell$ positive? Over which intervals is it negative? Over which, if any, is it zero?
b. Over which intervals is $\ell$ increasing? Over which is it decreasing? Over which, if any, is it constant?
c. What is the average value of $\ell$?

163. The graph of $y = \int_0^x \ell(t)dt,$ where $\ell$ is a piecewise linear function, is shown here.




a. Over which intervals is $\ell$ positive? Over which intervals is it negative? Over which, if any, is it zero?
b. Over which intervals is $\ell$ increasing? Over which is it decreasing? Over which intervals, if any, is it constant?

c. What is the average value of $\ell$?

In the following exercises, use a calculator to estimate the area under the curve by computing $T_{10}$, the average of the left- and right-endpoint Riemann sums using $N = 10$ rectangles. Then, using the Fundamental Theorem of Calculus, Part 2, determine the exact area.

164. **[T]** $y = x^2$ over $[0, 4]$

165. **[T]** $y = x^3 + 6x^2 + x - 5$ over $[-4, 2]$

166. **[T]** $y = \sqrt{x^3}$ over $[0, 6]$

167. **[T]** $y = \sqrt{x} + x^2$ over $[1, 9]$

168. **[T]** $\int (\cos x - \sin x) dx$ over $[0, \pi]$

169. **[T]** $\int \frac{4}{x^2} dx$ over $[1, 4]$

In the following exercises, evaluate each definite integral using the Fundamental Theorem of Calculus, Part 2.

170. $\int_{-1}^{2} \left(x^2 - 3x\right) dx$

171. $\int_{-2}^{3} \left(x^2 + 3x - 5\right) dx$

172. $\int_{-2}^{3} (t + 2)(t - 3)\, dt$

173. $\int_{2}^{3} \left(t^2 - 9\right)\left(4 - t^2\right) dt$

174. $\int_{1}^{2} x^9 dx$

175. $\int_{0}^{1} x^{99} dx$

176. $\int_{4}^{8} \left(4t^{5/2} - 3t^{3/2}\right) dt$

177. $\int_{1/4}^{4} \left(x^2 - \frac{1}{x^2}\right) dx$

178. $\int_{1}^{2} \frac{2}{x^3} dx$

179. $\int_{1}^{4} \frac{1}{2\sqrt{x}} dx$

180. $\int_{1}^{4} \frac{2 - \sqrt{t}}{t^2} dt$

181. $\int_{1}^{16} \frac{dt}{t^{1/4}}$

182. $\int_{0}^{2\pi} \cos\theta d\theta$

183. $\int_{0}^{\pi/2} \sin\theta d\theta$

184. $\int_{0}^{\pi/4} \sec^2\theta d\theta$

185. $\int_{0}^{\pi/4} \sec\theta \tan d\theta$

186. $\int_{\pi/3}^{\pi/4} \csc\theta \cot\theta d\theta$

187. $\int_{\pi/4}^{\pi/2} \csc^2\theta d\theta$

188. $\int_{1}^{2} \left(\frac{1}{t^2} - \frac{1}{t^3}\right) dt$

189. $\int_{-2}^{-1} \left(\frac{1}{t^2} - \frac{1}{t^3}\right) dt$

In the following exercises, use the evaluation theorem to express the integral as a function $F(x)$.

190. $\int_{a}^{x} t^2 dt$

191. $\int_1^x e^t \, dt$

192. $\int_0^x \cos t \, dt$

193. $\int_{-x}^x \sin t \, dt$

In the following exercises, identify the roots of the integrand to remove absolute values, then evaluate using the Fundamental Theorem of Calculus, Part 2.

194. $\int_{-2}^3 |x| \, dx$

195. $\int_{-2}^4 \left| t^2 - 2t - 3 \right| \, dt$

196. $\int_0^\pi |\cos t| \, dt$

197. $\int_{-\pi/2}^{\pi/2} |\sin t| \, dt$

198. Suppose that the number of hours of daylight on a given day in Seattle is modeled by the function $-3.75 \cos\left(\frac{\pi t}{6}\right) + 12.25$, with $t$ given in months and $t = 0$ corresponding to the winter solstice.

a. What is the average number of daylight hours in a year?
b. At which times $t_1$ and $t_2$, where $0 \leq t_1 < t_2 < 12$, do the number of daylight hours equal the average number?
c. Write an integral that expresses the total number of daylight hours in Seattle between $t_1$ and $t_2$.
d. Compute the mean hours of daylight in Seattle between $t_1$ and $t_2$, where $0 \leq t_1 < t_2 < 12$, and then between $t_2$ and $t_1$, and show that the average of the two is equal to the average day length.

199. Suppose the rate of gasoline consumption over the course of a year in the United States can be modeled by a sinusoidal function of the form $\left(11.21 - \cos\left(\frac{\pi t}{6}\right)\right) \times 10^9$ gal/mo.

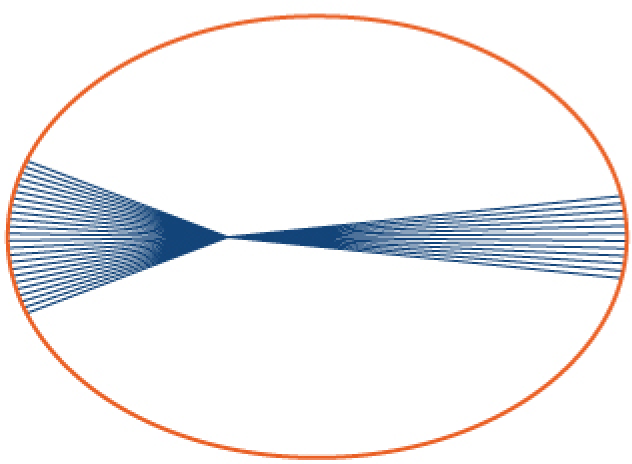
a. What is the average monthly consumption, and for which values of $t$ is the rate at time $t$ equal to the average rate?
b. What is the number of gallons of gasoline consumed in the United States in a year?

c. Write an integral that expresses the average monthly U.S. gas consumption during the part of the year between the beginning of April $(t = 3)$ and the end of September $(t = 9)$.

200. Explain why, if *f* is continuous over $[a, b]$, there is at least one point $c \in [a, b]$ such that $f(c) = \frac{1}{b-a} \int_a^b f(t)\, dt$.

201. Explain why, if *f* is continuous over $[a, b]$ and is not equal to a constant, there is at least one point $M \in [a, b]$ such that $f(M) = \frac{1}{b-a} \int_a^b f(t)\, dt$ and at least one point $m \in [a, b]$ such that $f(m) < \frac{1}{b-a} \int_a^b f(t)\, dt$.

202. Kepler's first law states that the planets move in elliptical orbits with the Sun at one focus. The closest point of a planetary orbit to the Sun is called the *perihelion* (for Earth, it currently occurs around January 3) and the farthest point is called the *aphelion* (for Earth, it currently occurs around July 4). Kepler's second law states that planets sweep out equal areas of their elliptical orbits in equal times. Thus, the two arcs indicated in the following figure are swept out in equal times. At what time of year is Earth moving fastest in its orbit? When is it moving slowest?



203. A point on an ellipse with major axis length 2*a* and minor axis length 2*b* has the coordinates $(a \cos \theta, b \sin \theta), 0 \le \theta \le 2\pi$.

a. Show that the distance from this point to the focus at $(-c, 0)$ is $d(\theta) = a + c \cos \theta$, where $c = \sqrt{a^2 - b^2}$.

b. Use these coordinates to show that the average distance $\bar{d}$ from a point on the ellipse to the focus at $(-c, 0)$, with respect to angle $\theta$, is *a*.

204. As implied earlier, according to Kepler's laws, Earth's orbit is an ellipse with the Sun at one focus. The perihelion for Earth's orbit around the Sun is 147,098,290 km and the aphelion is 152,098,232 km.

a. By placing the major axis along the *x*-axis, find the average distance from Earth to the Sun.
b. The classic definition of an astronomical unit (AU) is the distance from Earth to the Sun, and its value was computed as the average of the perihelion and aphelion distances. Is this definition justified?

205. The force of gravitational attraction between the Sun and a planet is $F(\theta) = \frac{GmM}{r^2(\theta)}$, where *m* is the mass of the planet, *M* is the mass of the Sun, *G* is a universal constant, and $r(\theta)$ is the distance between the Sun and the planet when the planet is at an angle $\theta$ with the major axis of its orbit. Assuming that *M*, *m*, and the ellipse parameters *a* and *b* (half-lengths of the major and minor axes) are given, set up—but do not evaluate—an integral that expresses in terms of $G, m, M, a, b$ the average gravitational force between the Sun and the planet.

206. The displacement from rest of a mass attached to a spring satisfies the simple harmonic motion equation $x(t) = A\cos(\omega t - \phi)$, where $\phi$ is a phase constant, $\omega$ is the angular frequency, and *A* is the amplitude. Find the average velocity, the average speed (magnitude of velocity), the average displacement, and the average distance from rest (magnitude of displacement) of the mass.

## Learning Objectives

1.4.1. Apply the basic integration formulas.
1.4.2. Explain the significance of the net change theorem.
1.4.3. Use the net change theorem to solve applied problems.
1.4.4. Apply the integrals of odd and even functions.

In this section, we use some basic integration formulas studied previously to solve some key applied problems. It is important to note that these formulas are presented in terms of *indefinite* integrals. Although definite and indefinite integrals are closely related, there are some key differences to keep in mind. A definite integral is either a number (when the limits of integration are constants) or a single function (when one or both of the limits of integration are variables). An indefinite integral represents a family of functions, all of which differ by a constant. As you become more familiar with integration, you will get a feel for when to use definite integrals and when to use indefinite integrals. You will naturally select the correct approach for a given problem without thinking too much about it. However, until these concepts are cemented in your mind, think carefully about whether you need a definite integral or an indefinite integral and make sure you are using the proper notation based on your choice.

## Basic Integration Formulas

Recall the integration formulas given in the table in Antiderivatives and the rule on properties of definite integrals. Let's look at a few examples of how to apply these rules.

**EXAMPLE 1.23**

### Integrating a Function Using the Power Rule

Use the power rule to integrate the function $\int_1^4 \sqrt{t}\,(1+t)\,dt$.

**[Show Solution]**

**CHECKPOINT 1.21**

Find the definite integral of $f(x) = x^2 - 3x$ over the interval $[1, 3]$.

## The Net Change Theorem

The **net change theorem** considers the integral of a *rate of change*. It says that when a quantity changes, the new value equals the initial value plus the integral of the rate of change of that quantity. The formula can be expressed in two ways. The second is more familiar; it is simply the definite integral.

**THEOREM 1.6**

**Net Change Theorem**

The new value of a changing quantity equals the initial value plus the integral of the rate of change:

$$F(b) = F(a) + \int_a^b F'(x)\,dx \tag{1.18}$$

or

$$\int_a^b F'(x)\,dx = F(b) - F(a).$$

Subtracting $F(a)$ from both sides of the first equation yields the second equation. Since they are equivalent formulas, which one we use depends on the application.

The significance of the net change theorem lies in the results. Net change can be applied to area, distance, and volume, to name only a few applications. Net change accounts for negative quantities automatically without having to write more than one integral. To illustrate, let's apply the net change theorem to a velocity function in which the result is displacement.

We looked at a simple example of this in The Definite Integral. Suppose a car is moving due north (the positive direction) at 40 mph between 2 p.m. and 4 p.m., then the car moves south at 30 mph between 4 p.m. and 5 p.m. We can graph this motion as shown in Figure 1.32.

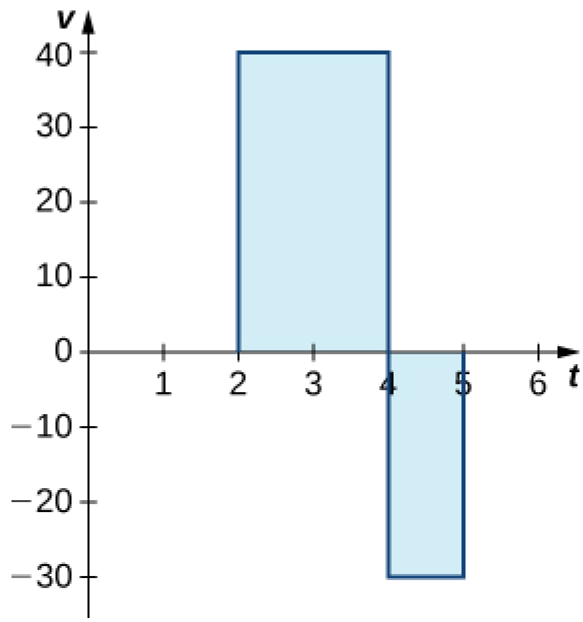


**Figure 1.32** The graph shows speed versus time for the given motion of a car.

Just as we did before, we can use definite integrals to calculate the net displacement as well as the total distance traveled. The net displacement is given by

$$\begin{aligned}\int_2^5 v(t)dt &= \int_2^4 40dt + \int_4^5 -30dt \\ &= 80 - 30 \\ &= 50.\end{aligned}$$

Thus, at 5 p.m. the car is 50 mi north of its starting position. The total distance traveled is given by

$$\begin{aligned}\int_2^5 |v(t)|\, dt &= \int_2^4 40dt + \int_4^5 30dt \\ &= 80 + 30 \\ &= 110.\end{aligned}$$

Therefore, between 2 p.m. and 5 p.m., the car traveled a total of 110 mi.

To summarize, net displacement may include both positive and negative values. In other words, the velocity function accounts for both forward distance and backward distance.

To find net displacement, integrate the velocity function over the interval. Total distance traveled, on the other hand, is always positive. To find the total distance traveled by an object, regardless of direction, we need to integrate the absolute value of the velocity function.

**EXAMPLE 1.24**

### Finding Net Displacement

Given a velocity function $v(t) = 3t - 5$ (in meters per second) for a particle in motion from time $t = 0$ to time $t = 3$, find the net displacement of the particle.

[Show Solution]

**EXAMPLE 1.25**

### Finding the Total Distance Traveled

Use Example 1.24 to find the total distance traveled by a particle according to the velocity function $v(t) = 3t - 5$ m/sec over a time interval $[0, 3]$.

[Show Solution]

**CHECKPOINT 1.22**

Find the net displacement and total distance traveled in meters given the velocity function $f(t) = \frac{1}{2}e^t - 2$ over the interval $[0, 2]$.

## Applying the Net Change Theorem

The net change theorem can be applied to the flow and consumption of fluids, as shown in Example 1.26.

### EXAMPLE 1.26

#### How Many Gallons of Gasoline Are Consumed?

If the motor on a motorboat is started at $t = 0$ and the boat consumes gasoline at $5 - t^3$ gal/hr for the first hour, how much gasoline is used in the first hour?

[Show Solution]

### EXAMPLE 1.27

#### Chapter Opener: Iceboats

**Figure 1.34** (credit: modification of work by Carter Brown, Flickr)

As we saw at the beginning of the chapter, top iceboat racers (Figure 1.1) can attain speeds of up to five times the wind speed. Andrew is an intermediate iceboater, though, so he attains speeds equal to only twice the wind speed. Suppose Andrew takes his iceboat out one morning when a light 5-mph breeze has been blowing all morning. As Andrew gets his iceboat set up, though, the wind begins to pick up. During his first half hour of iceboating, the wind speed increases according to the function $v(t) = 20t + 5$. For the second half hour of Andrew's outing, the wind remains steady at 15 mph. In other words, the wind speed is given by

$$v(t) = \begin{cases} 20t + 5 & \text{for} \quad 0 \le t \le \frac{1}{2} \\ 15 & \text{for} \quad \frac{1}{2} \le t \le 1. \end{cases}$$

Recalling that Andrew's iceboat travels at twice the wind speed, and assuming he moves in a straight line away from his starting point, how far is Andrew from his starting point after 1 hour?

**[Show Solution]**

**CHECKPOINT 1.23**

Suppose that, instead of remaining steady during the second half hour of Andrew's outing, the wind starts to die down according to the function $v(t) = -10t + 15$. In other words, the wind speed is given by

$$v(t) = \begin{cases} 20t + 5 & \text{for} \quad 0 \le t \le \frac{1}{2} \\ -10t + 15 & \text{for} \quad \frac{1}{2} \le t \le 1. \end{cases}$$

Under these conditions, how far from his starting point is Andrew after 1 hour?

## Integrating Even and Odd Functions

We saw in Functions and Graphs that an even function is a function in which $f(-x) = f(x)$ for all $x$ in the domain—that is, the graph of the curve is unchanged when $x$ is replaced with $-x$. The graphs of even functions are symmetric about the $y$-axis. An odd function is one in which $f(-x) = -f(x)$ for all $x$ in the domain, and the graph of the function is symmetric about the origin.

Integrals of even functions, when the limits of integration are from $-a$ to $a$, involve two equal areas, because they are symmetric about the $y$-axis. Integrals of odd functions, when the limits of integration are similarly $[-a, a]$, evaluate to zero because the areas above and below the $x$-axis are equal.

### RULE: INTEGRALS OF EVEN AND ODD FUNCTIONS

For continuous even functions such that $f(-x) = f(x)$,

$$\int_{-a}^{a} f(x)\,dx = 2\int_{0}^{a} f(x)\,dx.$$

For continuous odd functions such that $f(-x) = -f(x)$,

$$\int_{-a}^{a} f(x)\,dx = 0.$$

### EXAMPLE 1.28

#### Integrating an Even Function

Integrate the even function $\int_{-2}^{2} \left(3x^8 - 2\right) dx$ and verify that the integration formula for even functions holds.

**[Show Solution]**

### EXAMPLE 1.29

### Integrating an Odd Function

Evaluate the definite integral of the odd function $-5\sin x$ over the interval $[-\pi, \pi]$ .

[Show Solution]

**CHECKPOINT 1.24**

Integrate the function $\int_{-2}^{2} x^4 dx.$

## Section 1.4 Exercises

Use basic integration formulas to compute the following antiderivatives of definite integrals or definite integrals.

207. $\int \left(\sqrt{x} - \frac{1}{\sqrt{x}}\right) dx$

208. $\int \left(e^{2x} - \frac{1}{2}e^{x/2}\right) dx$

209. $\int \frac{dx}{2x}$

210. $\int \frac{x-1}{x^2} dx$

211. $\int_0^{\pi} (\sin x - \cos x)\, dx$

212. $\int_0^{\pi/2} (x - \sin x)\, dx$

213. Write an integral that expresses the increase in the perimeter $P(s)$ of a square when its side length *s* increases from 2 units to 4 units and evaluate the integral.

214. Write an integral that quantifies the change in the area $A(s) = s^2$ of a square when the side length doubles from *S* units to 2*S* units and evaluate the integral.

215. A regular *N*-gon (an *N*-sided polygon with sides that have equal length *s*, such as a pentagon or hexagon) has perimeter *Ns*. Write an integral that expresses the increase in perimeter of a regular *N*-gon when the length of each side increases from 1 unit to 2 units and evaluate the integral.

216. The area of a regular pentagon with side length $a > 0$ is $pa^2$ with $p = \frac{1}{4}\sqrt{5 + \sqrt{5 + 2\sqrt{5}}}$. The Pentagon in Washington, DC, has inner sides of length 360 ft and outer sides of length 920 ft. Write an integral to express the area of the roof of the Pentagon according to these dimensions and evaluate this area.

217. A dodecahedron is a Platonic solid with a surface that consists of 12 pentagons, each of equal area. By how much does the surface area of a dodecahedron increase as the side length of each pentagon doubles from 1 unit to 2 units?

218. An icosahedron is a Platonic solid with a surface that consists of 20 equilateral triangles. By how much does the surface area of an icosahedron increase as the side length of each triangle doubles from *a* unit to 2*a* units?

219. Write an integral that quantifies the change in the area of the surface of a cube when its side length doubles from *s* unit to 2*s* units and evaluate the integral.

220. Write an integral that quantifies the increase in the volume of a cube when the side length doubles from *s* unit to 2*s* units and evaluate the integral.

221. Write an integral that quantifies the increase in the surface area of a sphere as its radius doubles from *R* unit to 2*R* units and evaluate the integral.

222. Write an integral that quantifies the increase in the volume of a sphere as its radius doubles from *R* unit to 2*R* units and evaluate the integral.

223. Suppose that a particle moves along a straight line with velocity $v(t) = 4 - 2t$, where $0 \leq t \leq 2$ (in meters per second). Find the displacement at time *t* and the total distance traveled up to $t = 2$.

224. Suppose that a particle moves along a straight line with velocity defined by $v(t) = t^2 - 3t - 18$, where $0 \leq t \leq 6$ (in meters per second). Find the displacement at time *t* and the total distance traveled up to $t = 6$.

225. Suppose that a particle moves along a straight line with velocity defined by $v(t) = |2t - 6|$, where $0 \leq t \leq 6$ (in meters per second). Find the displacement at time *t* and the total distance traveled up to $t = 6$.

226. Suppose that a particle moves along a straight line with acceleration defined by $a(t) = t - 3$, where $0 \leq t \leq 6$ (in meters per second). Find the velocity and displacement at time *t* and the total distance traveled up to $t = 6$ if $v(0) = 3$ and $d(0) = 0$.

227. A ball is thrown upward from a height of 1.5 m at an initial speed of 40 m/sec. Acceleration resulting from gravity is –9.8 m/sec$^2$. Neglecting air resistance, solve for the velocity $v(t)$ and the height $h(t)$ of the ball *t* seconds after it is thrown and before it returns to the ground.

228. A ball is thrown upward from a height of 3 m at an initial speed of 60 m/sec. Acceleration resulting from gravity is –9.8 m/sec$^2$. Neglecting air resistance, solve for the velocity $v(t)$ and the height $h(t)$ of the ball *t* seconds after it is thrown and before it returns to the ground.

229. The area $A(t)$ of a circular shape is growing at a constant rate. If the area increases from $4\pi$ units to $9\pi$ units between times $t = 2$ and $t = 3$, find the net change in the radius during that time.

230. A spherical balloon is being inflated at a constant rate. If the volume of the balloon changes from $36\pi$ in.$^3$ to $288\pi$ in.$^3$ between time $t = 30$ and $t = 60$ seconds, find the net change in the radius of the balloon during that time.

231. Water flows into a conical tank with cross-sectional area $\pi x^2$ at height *x* and volume $\frac{\pi x^3}{3}$ up to height *x*. If water flows into the tank at a rate of 1 m$^3$/min, find the height of water in the tank after 5 min. Find the change in height between 5 min and 10 min.

232. A horizontal cylindrical tank has cross-sectional area $A(x) = 4\left(6x - x^2\right) m^2$ at height *x* meters above the bottom when $x \leq 3$.

a. The volume *V* between heights *a* and *b* is $\int_a^b A(x)\,dx$. Find the volume at heights between 2 m and 3 m.
b. Suppose that oil is being pumped into the tank at a rate of 50 L/min. Using the chain rule, $\frac{dx}{dt} = \frac{dx}{dV}\frac{dV}{dt}$, at how many meters per minute is the height of oil in the tank changing, expressed in terms of *x*, when the height is at *x* meters?
c. How long does it take to fill the tank to 3 m starting from a fill level of 2 m?

233. The following table lists the electrical power in gigawatts—the rate at which energy is consumed—used in a certain city for different hours of the day, in a typical 24-hour period, with hour 1 corresponding to midnight to 1 a.m.

| Hour | Power | Hour | Power |
|---|---|---|---|
| 1 | 28 | 13 | 48 |
| 2 | 25 | 14 | 49 |
| 3 | 24 | 15 | 49 |
| 4 | 23 | 16 | 50 |
| 5 | 24 | 17 | 50 |
| 6 | 27 | 18 | 50 |
| 7 | 29 | 19 | 46 |
| 8 | 32 | 20 | 43 |
| 9 | 34 | 21 | 42 |
| 10 | 39 | 22 | 40 |
| 11 | 42 | 23 | 37 |
| 12 | 46 | 24 | 34 |

Find the total amount of energy in gigawatt-hours (gW-h) consumed by the city in a typical 24-hour period.

234. The average residential electrical power use (in hundreds of watts) per hour is given in the following table.

| Hour | Power | Hour | Power |
|---|---|---|---|
| 1 | 8 | 13 | 12 |
| 2 | 6 | 14 | 13 |
| 3 | 5 | 15 | 14 |
| 4 | 4 | 16 | 15 |
| 5 | 5 | 17 | 17 |
| 6 | 6 | 18 | 19 |

| Hour | Power | Hour | Power |
|---|---|---|---|
| 7 | 7 | 19 | 18 |
| 8 | 8 | 20 | 17 |
| 9 | 9 | 21 | 16 |
| 10 | 10 | 22 | 16 |
| 11 | 10 | 23 | 13 |
| 12 | 11 | 24 | 11 |

a. Compute the average total energy used in a day in kilowatt-hours (kWh).
b. If a ton of coal generates 1842 kWh, how long does it take for an average residence to burn a ton of coal?
c. Explain why the data might fit a plot of the form $p(t) = 11.5 - 7.5\sin\left(\frac{\pi t}{12}\right)$.

235. The data in the following table are used to estimate the average power output produced by Peter Sagan for each of the last 18 sec of Stage 1 of the 2012 Tour de France.

| Second | Watts | Second | Watts |
|---|---|---|---|
| 1 | 600 | 10 | 1200 |
| 2 | 500 | 11 | 1170 |
| 3 | 575 | 12 | 1125 |
| 4 | 1050 | 13 | 1100 |
| 5 | 925 | 14 | 1075 |
| 6 | 950 | 15 | 1000 |
| 7 | 1050 | 16 | 950 |
| 8 | 950 | 17 | 900 |
| 9 | 1100 | 18 | 780 |

**Table 1.6** Average Power Output *Source*: sportsexercisengineering.com

Estimate the net energy used in kilojoules (kJ), noting that 1W = 1 J/s, and the average power output by Sagan during this time interval.

236. The data in the following table are used to estimate the average power output produced by Peter Sagan for each 15-min interval of Stage 1 of the 2012 Tour de France.

| Minutes | Watts | Minutes | Watts |
|---|---|---|---|
| 15 | 200 | 165 | 170 |
| 30 | 180 | 180 | 220 |
| 45 | 190 | 195 | 140 |
| 60 | 230 | 210 | 225 |
| 75 | 240 | 225 | 170 |
| 90 | 210 | 240 | 210 |
| 105 | 210 | 255 | 200 |
| 120 | 220 | 270 | 220 |
| 135 | 210 | 285 | 250 |
| 150 | 150 | 300 | 400 |

**Table 1.7** Average Power Output *Source*: sportsexercisengineering.com

Estimate the net energy used in kilojoules, noting that 1W = 1 J/s.

237. The distribution of incomes as of 2012 in the United States in $5000 increments is given in the following table. The *k*th row denotes the percentage of households with incomes between $\$5000xk$ and $5000xk + 4999$. The row $k = 40$ contains all households with income between $200,000 and $250,000 and $k = 41$ accounts for all households with income exceeding $250,000.

| | | | |
|---|---|---|---|
| 0 | 3.5 | 21 | 1.5 |
| 1 | 4.1 | 22 | 1.4 |
| 2 | 5.9 | 23 | 1.3 |
| | | | |

| | | | |
|---|---|---|---|
| 3 | 5.7 | 24 | 1.3 |
| 4 | 5.9 | 25 | 1.1 |
| 5 | 5.4 | 26 | 1.0 |
| 6 | 5.5 | 27 | 0.75 |
| 7 | 5.1 | 28 | 0.8 |
| 8 | 4.8 | 29 | 1.0 |
| 9 | 4.1 | 30 | 0.6 |
| 10 | 4.3 | 31 | 0.6 |
| 11 | 3.5 | 32 | 0.5 |
| 12 | 3.7 | 33 | 0.5 |
| 13 | 3.2 | 34 | 0.4 |
| 14 | 3.0 | 35 | 0.3 |
| 15 | 2.8 | 36 | 0.3 |
| 16 | 2.5 | 37 | 0.3 |
| 17 | 2.2 | 38 | 0.2 |
| 18 | 2.2 | 39 | 1.8 |
| 19 | 1.8 | 40 | 2.3 |
| 20 | 2.1 | 41 | |

**Table 1.8** Income Distributions *Source*: http://www.census.gov/prod/2013pubs/p60-245.pdf

a. Estimate the percentage of U.S. households in 2012 with incomes less than $55,000.
b. What percentage of households had incomes exceeding $85,000?
c. Plot the data and try to fit its shape to that of a graph of the form $a\,(x+c)\,e^{-b(x+e)}$ for suitable $a, b, c$.

238. Newton's law of gravity states that the gravitational force exerted by an object of mass *M* and one of mass *m* with centers that are separated by a distance *r* is $F = G\frac{mM}{r^2}$, with *G* an empirical constant $G = 6.67x10^{-11}\ m^3/\left(kg \cdot s^2\right)$. The work done by a variable force over an interval $[a, b]$ is defined as $W = \int_a^b F(x)\,dx.$ If Earth has mass $5.97219 \times 10^{24}$ and radius 6371 km, compute the amount of work to elevate a polar weather satellite of mass 1400 kg to its orbiting altitude of 850 km above Earth.

239. For a given motor vehicle, the maximum achievable deceleration from braking is approximately 7 m/sec$^2$ on dry concrete. On wet asphalt, it is approximately 2.5 m/sec$^2$. Given that 1 mph corresponds to 0.447 m/sec, find the total distance that a car travels in meters on dry concrete after the brakes are applied until it comes to a complete stop if the initial velocity is 67 mph (30 m/sec) or if the initial braking velocity is 56 mph (25 m/sec). Find the corresponding distances if the surface is slippery wet asphalt.

240. John is a 25-year old man who weighs 160 lb. He burns $500 - 50t$ calories/hr while riding his bike for *t* hours. If an oatmeal cookie has 55 cal and John eats 4*t* cookies during the *t*th hour, how many net calories has he lost after 3 hours riding his bike?

241. Sandra is a 25-year old woman who weighs 120 lb. She burns $300 - 50t$ cal/hr while walking on her treadmill. Her caloric intake from drinking Gatorade is 100*t* calories during the *t*th hour. What is her net decrease in calories after walking for 3 hours?

242. A motor vehicle has a maximum efficiency of 33 mpg at a cruising speed of 40 mph. The efficiency drops at a rate of 0.1 mpg/mph between 40 mph and 50 mph, and at a rate of 0.4 mpg/mph between 50 mph and 80 mph. What is the efficiency in miles per gallon if the car is cruising at 50 mph? What is the efficiency in miles per gallon if the car is cruising at 80 mph? If gasoline costs $3.50/gal, what is the cost of fuel to drive 50 mi at 40 mph, at 50 mph, and at 80 mph?

243. Although some engines are more efficient at given a horsepower than others, on average, fuel efficiency decreases with horsepower at a rate of $1/25$ mpg/horsepower. If a typical 50-horsepower engine has an average fuel efficiency of 32 mpg, what is the average fuel efficiency of an engine with the following horsepower: 150, 300, 450?

244. **[T]** The following table lists the 2013 schedule of federal income tax versus taxable income.

| Taxable Income Range | The Tax Is ... | ... Of the Amount Over |
|---|---|---|
| \$0–\$8925 | 10% | \$0 |
| \$8925–\$36,250 | \$892.50 + 15% | \$8925 |
| \$36,250–\$87,850 | \$4,991.25 + 25% | \$36,250 |
| \$87,850–\$183,250 | \$17,891.25 + 28% | \$87,850 |
| \$183,250–\$398,350 | \$44,603.25 + 33% | \$183,250 |
| \$398,350–\$400,000 | \$115,586.25 + 35% | \$398,350 |
| > \$400,000 | \$116,163.75 + 39.6% | \$400,000 |

**Table 1.9** Federal Income Tax Versus Taxable Income *Source*: http://www.irs.gov/pub/irs-prior/i1040tt--2013.pdf.

Suppose that Steve just received a \$10,000 raise. How much of this raise is left after federal taxes if Steve's salary before receiving the raise was \$40,000? If it was \$90,000? If it was \$385,000?

245. **[T]** The following table provides hypothetical data regarding the level of service for a certain highway.

| Highway Speed Range (mph) | Vehicles per Hour per Lane | Density Range (vehicles/mi) |
|---|---|---|
| > 60 | < 600 | < 10 |
| 60–57 | 600–1000 | 10–20 |
| 57–54 | 1000–1500 | 20–30 |
| 54–46 | 1500–1900 | 30–45 |
| 46–30 | 1900–2100 | 45–70 |
| <30 | Unstable | 70–200 |

**Table 1.10**

a. Plot vehicles per hour per lane on the *x*-axis and highway speed on the *y*-axis.

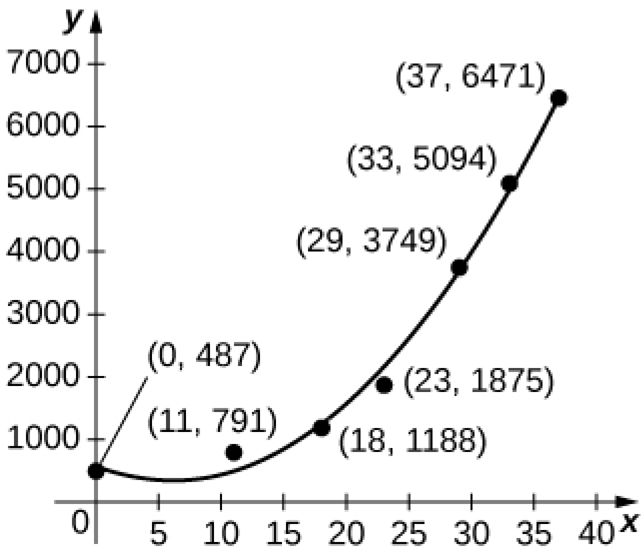
b. Compute the average decrease in speed (in miles per hour) per unit increase in congestion (vehicles per hour per lane) as the latter increases from 600 to 1000, from 1000 to 1500, and from 1500 to 2100. Does the decrease in miles per hour depend linearly on the increase in vehicles per hour per lane?
c. Plot minutes per mile (60 times the reciprocal of miles per hour) as a function of vehicles per hour per lane. Is this function linear?

For the next two exercises use the data in the following table, which displays bald eagle populations from 1963 to 2000 in the continental United States.

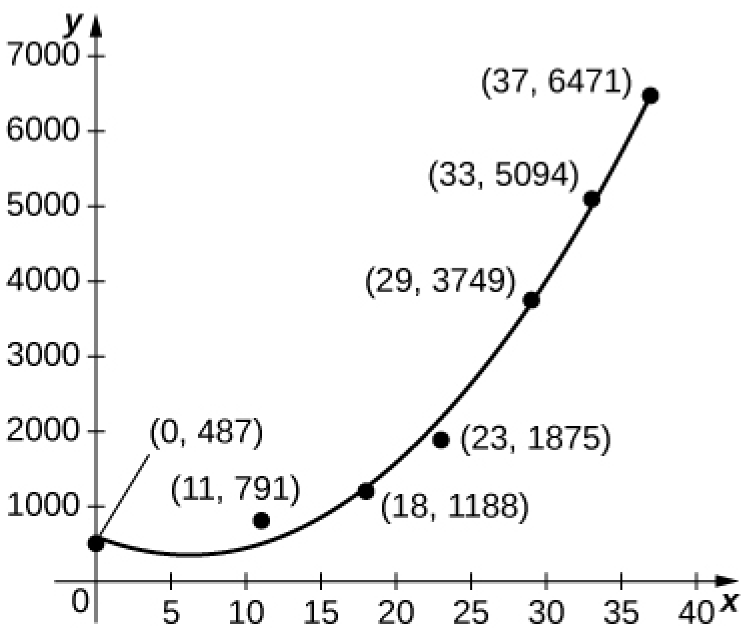
| Year | Population of Breeding Pairs of Bald Eagles |
|---|---|
| 1963 | 487 |
| 1974 | 791 |
| 1981 | 1188 |
| 1986 | 1875 |
| 1992 | 3749 |
| 1996 | 5094 |
| 2000 | 6471 |

**Table 1.11** Population of Breeding Bald Eagle Pairs *Source*: http://www.fws.gov/Midwest/eagle/population/chtofprs.html.

246. **[T]** The graph below plots the quadratic $p(t) = 6.48t^2 - 80.3\,1t + 585.69$ against the data in preceding table, normalized so that $t = 0$ corresponds to 1963. Estimate the average number of bald eagles per year present for the 37 years by computing the average value of $p$ over $[0, 37]$ .

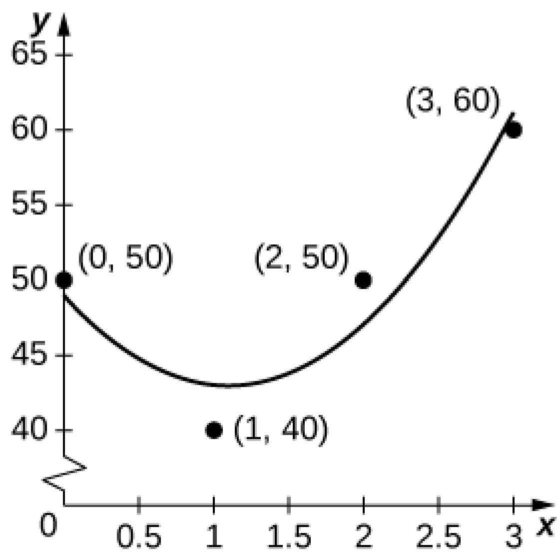



247. **[T]** The graph below plots the cubic $p(t) = 0.07t^3 + 2.42t^2 - 25.63t + 521.23$ against the data in the preceding table, normalized so that $t = 0$ corresponds to 1963. Estimate the average number of bald eagles per year present for the 37 years by computing the average value of *p* over $[0, 37]$.




248. **[T]** Suppose you go on a road trip and record your speed at every half hour, as compiled in the following table. The best quadratic fit to the data is $q(t) = 5x^2 - 11x + 49$, shown in the accompanying graph. Integrate *q* to estimate the total distance driven over the 3 hours.

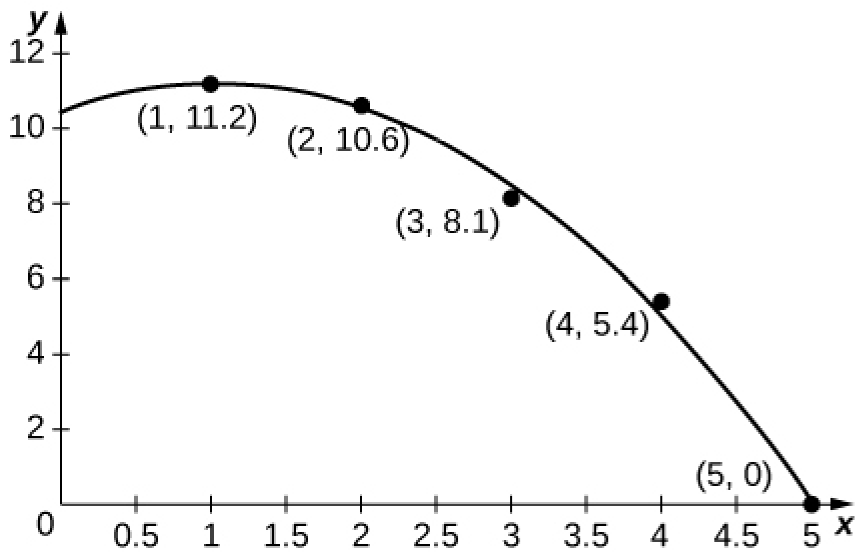
| Time (hr) | Speed (mph) |
|---|---|
| 0 (start) | 50 |
| 1 | 40 |
| 2 | 50 |
| 3 | 60 |




As a car accelerates, it does not accelerate at a constant rate; rather, the acceleration is variable. For the following exercises, use the following table, which contains the acceleration measured at every second as a driver merges onto a freeway.

| Time (sec) | Acceleration (mph/sec) |
|---|---|
| 1 | 11.2 |
| 2 | 10.6 |
| 3 | 8.1 |
| 4 | 5.4 |
| 5 | 0 |

249. **[T]** The accompanying graph plots the best quadratic fit, $a(t) = -0.70t^2 + 1.44t + 10.44$, to the data from the preceding table. Compute the average value of $a(t)$ to estimate the average acceleration between $t = 0$ and $t = 5$.



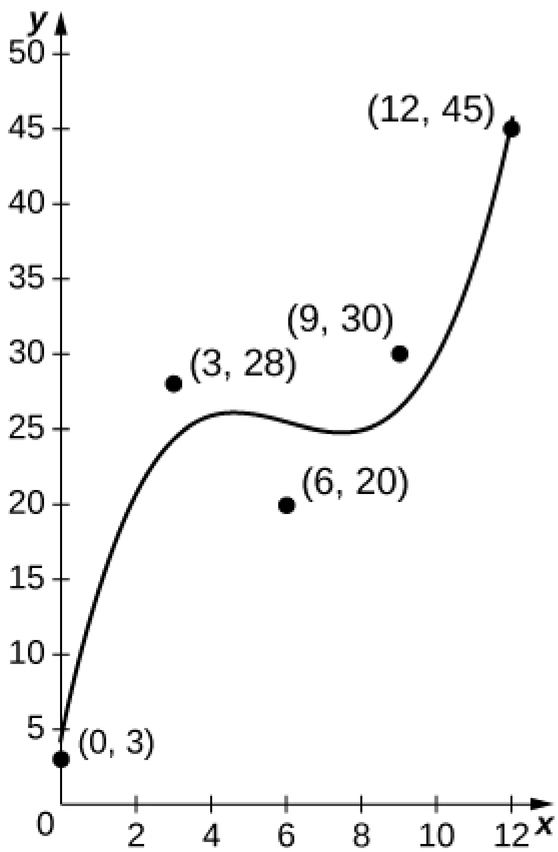

250. **[T]** Using your acceleration equation from the previous exercise, find the corresponding velocity equation. Assuming the final velocity is 0 mph, find the velocity at time $t = 0$.

251. **[T]** Using your velocity equation from the previous exercise, find the corresponding distance equation, assuming your initial distance is 0 mi. How far did you travel while you accelerated your car? (*Hint:* You will need to convert time units.)

252. **[T]** The number of hamburgers sold at a restaurant throughout the day is given in the following table, with the accompanying graph plotting the best cubic fit to the data, $b(t) = 0.12t^3 - 2.13t^3 + 12.13t + 3.91$, with $t = 0$ corresponding to 9 a.m. and $t = 12$ corresponding to 9 p.m. Compute the average value of $b(t)$ to estimate the average number of hamburgers sold per hour.

| Hours Past Midnight | No. of Burgers Sold |
|---|---|
| 9 | 3 |
| 12 | 28 |
| 15 | 20 |
| 18 | 30 |

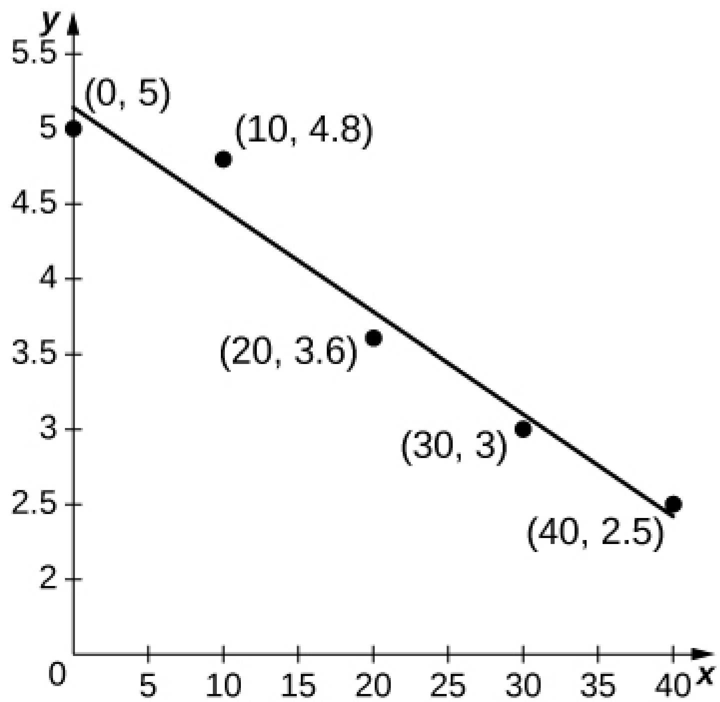
| Hours Past Midnight | No. of Burgers Sold |
|---|---|
| 21 | 45 |




253. **[T]** An athlete runs by a motion detector, which records her speed, as displayed in the following table. The best linear fit to this data, $\ell(t) = -0.068t + 5.14$, is shown in the accompanying graph. Use the average value of $\ell(t)$ between $t = 0$ and $t = 40$ to estimate the runner's average speed.

| Minutes | Speed (m/sec) |
|---|---|
| 0 | 5 |
| 10 | 4.8 |
| 20 | 3.6 |
| 30 | 3.0 |

| Minutes | Speed (m/sec) |
|---|---|
| 40 | 2.5 |



## Learning Objectives

1.5.1. Use substitution to evaluate indefinite integrals.
1.5.2. Use substitution to evaluate definite integrals.

The Fundamental Theorem of Calculus gave us a method to evaluate integrals without using Riemann sums. The drawback of this method, though, is that we must be able to find an antiderivative, and this is not always easy. In this section we examine a technique, called **integration by substitution**, to help us find antiderivatives. Specifically, this method helps us find antiderivatives when the integrand is the result of a chain-rule derivative.

At first, the approach to the substitution procedure may not appear very obvious. However, it is primarily a visual task—that is, the integrand shows you what to do; it is a matter of recognizing the form of the function. So, what are we supposed to see? We are looking for an integrand of the form $f\left[g\left(x\right)\right]g'\left(x\right)dx.$ For example, in the integral $\int \left(x^2-3\right)^3 2xdx,$ we have $f(x)=x^3$, $g(x)=x^2-3$, and $g'(x)=2x$. Then,

$$f\left[g\left(x\right)\right]g'\left(x\right)=\left(x^2-3\right)^3\left(2x\right),$$

and we see that our integrand is in the correct form.

The method is called *substitution* because we substitute part of the integrand with the variable *u* and part of the integrand with *du*. It is also referred to as **change of variables** because we are changing variables to obtain an expression that is easier to work with for applying the integration rules.

### THEOREM 1.7

### Substitution with Indefinite Integrals

Let $u=g\left(x\right),$ , where $g'\left(x\right)$ is continuous over an interval, let $f\left(x\right)$ be continuous over the corresponding range of *g*, and let $F(x)$ be an antiderivative of $f(x)$. Then,

$$\begin{aligned}\int f\left[g\left(x\right)\right]g'\left(x\right)dx &= \int f\left(u\right)du \\ &= F\left(u\right)+C \\ &= F\left(g\left(x\right)\right)+C.\end{aligned} \tag{1.19}$$

## Proof

Let *f*, *g*, *u*, and *F* be as specified in the theorem. Then

$$\begin{aligned} \frac{d}{dx}F(g(x)) &= F'(g(x))g'(x) \\ &= f[g(x)]\,g'(x). \end{aligned}$$

Integrating both sides with respect to *x*, we see that

$$\int f[g(x)]\,g'(x)dx = F(g(x)) + C.$$

If we now substitute $u = g(x)$, and $du = g'(x)\,dx$, we get

$$\begin{aligned} \int f[g(x)]\,g'(x)dx &= \int f(u)\,du \\ &= F(u) + C \\ &= F(g(x)) + C. \end{aligned}$$

□

Returning to the problem we looked at originally, we let $u = x^2 - 3$ and then $du = 2xdx$. Rewrite the integral in terms of *u*:

$$\int \underbrace{\left(x^2 - 3\right)}_{u}^{3} \underbrace{(2xdx)}_{du} = \int u^3 du.$$

Using the power rule for integrals, we have

$$\int u^3 du = \frac{u^4}{4} + C.$$

Substitute the original expression for *x* back into the solution:

$$\frac{u^4}{4} + C = \frac{\left(x^2 - 3\right)^4}{4} + C.$$

We can generalize the procedure in the following Problem-Solving Strategy.

### PROBLEM-SOLVING STRATEGY: INTEGRATION BY SUBSTITUTION

1. Look carefully at the integrand and select an expression $g(x)$ within the integrand to set equal to *u*. Let's select $g(x)$. such that $g'(x)$ is also part of the integrand.
2. Substitute $u = g(x)$ and $du = g'(x)dx$. into the integral.
3. We should now be able to evaluate the integral with respect to *u*. If the integral can't be evaluated we need to go back and select a different expression to use as *u*.
4. Evaluate the integral in terms of *u*.
5. Write the result in terms of *x* and the expression $g(x)$.

### EXAMPLE 1.30

#### Using Substitution to Find an Antiderivative

Use substitution to find the antiderivative $\int 6x\left(3x^2+4\right)^4 dx.$

[Show Solution]

### CHECKPOINT 1.25

Use substitution to find the antiderivative $\int 3x^2\left(x^3-3\right)^2 dx.$

Sometimes we need to adjust the constants in our integral if they don't match up exactly with the expressions we are substituting.

### EXAMPLE 1.31

### Using Substitution with Alteration

Use substitution to find $\int z\sqrt{z^2 - 5}dz.$

[Show Solution]

**CHECKPOINT 1.26**

Use substitution to find $\int x^2\left(x^3 + 5\right)^9 dx.$

**EXAMPLE 1.32**

### Using Substitution with Integrals of Trigonometric Functions

Use substitution to evaluate the integral $\int \frac{\sin t}{\cos^3 t}dt.$

[Show Solution]

**CHECKPOINT 1.27**

Use substitution to evaluate the integral $\int \frac{\cos t}{\sin^2 t}dt.$

Sometimes we need to manipulate an integral in ways that are more complicated than just multiplying or dividing by a constant. We need to eliminate all the expressions within the integrand that are in terms of the original variable. When we are done, *u* should be the only variable in the integrand. In some cases, this means solving for the original variable in terms of *u*. This technique should become clear in the next example.

**EXAMPLE 1.33**

### Finding an Antiderivative Using *u*-Substitution

Use substitution to find the antiderivative $\int \frac{x}{\sqrt{x-1}} dx$.

[Show Solution]

**CHECKPOINT 1.28**

Use substitution to evaluate the indefinite integral $\int \cos^3 t \sin t \, dt$.

## Substitution for Definite Integrals

Substitution can be used with definite integrals, too. However, using substitution to evaluate a definite integral requires a change to the limits of integration. If we change variables in the integrand, the limits of integration change as well.

**THEOREM 1.8**

### Substitution with Definite Integrals

Let $u = g(x)$ and let $g'$ be continuous over an interval $[a, b]$, and let *f* be continuous over the range of $u = g(x)$. Then,

$$\int_a^b f(g(x))\, g'(x) dx = \int_{g(a)}^{g(b)} f(u)\, du.$$

Although we will not formally prove this theorem, we justify it with some calculations here. From the substitution rule for indefinite integrals, if $F(x)$ is an antiderivative of $f(x)$, we have

$$\int f(g(x))\, g'(x)dx = F(g(x)) + C.$$

Then

$$\begin{aligned}\int_a^b f[g(x)]\, g'(x)dx &= F(g(x))\big|_{x=a}^{x=b} \\ &= F(g(b)) - F(g(a)) \\ &= F(u)\big|_{u=g(a)}^{u=g(b)} \\ &= \int_{g(a)}^{g(b)} f(u)du,\end{aligned} \tag{1.20}$$

and we have the desired result.

**EXAMPLE 1.34**

### Using Substitution to Evaluate a Definite Integral

Use substitution to evaluate $\int_0^1 x^2\left(1+2x^3\right)^5 dx.$

[Show Solution]

**CHECKPOINT 1.29**

Use substitution to evaluate the definite integral $\int_{-1}^{0} y\left(2y^2-3\right)^5 dy.$

#### EXAMPLE 1.35

### Using Substitution with an Exponential Function

Use substitution to evaluate $\int_0^1 xe^{4x^2+3}\,dx.$

[Show Solution]

#### CHECKPOINT 1.30

Use substitution to evaluate $\int_0^1 x^2\cos\left(\frac{\pi}{2}x^3\right)dx.$

Substitution may be only one of the techniques needed to evaluate a definite integral. All of the properties and rules of integration apply independently, and trigonometric functions may need to be rewritten using a trigonometric identity before we can apply substitution. Also, we have the option of replacing the original expression for *u* after we find the antiderivative, which means that we do not have to change the limits of integration. These two approaches are shown in Example 1.36.

#### EXAMPLE 1.36

### Using Substitution to Evaluate a Trigonometric Integral

Use substitution to evaluate $\int_0^{\pi/2} \cos^2\theta\,d\theta.$

[Show Solution]

## Section 1.5 Exercises

254. Why is $u$-substitution referred to as *change of variable*?

255. 2. If $f = g \circ h$, when reversing the chain rule, $\frac{d}{dx}(g \circ h)(x) = g'(h(x))h'(x)$, should you take $u = g(x)$ or $u = h(x)$?

In the following exercises, verify each identity using differentiation. Then, using the indicated $u$-substitution, identify $f$ such that the integral takes the form $\int f(u)\,du$.

256. $\int x\sqrt{x+1}\,dx = \frac{2}{15}(x+1)^{3/2}(3x-2) + C; u = x+1$

257. For $x > 1$ : $\int \frac{x^2}{\sqrt{x-1}}dx = \frac{2}{15}\sqrt{x-1}\left(3x^2+4x+8\right) + C; u = x-1$

258. $\int x\sqrt{4x^2+9}\,dx = \frac{1}{12}\left(4x^2+9\right)^{3/2} + C; u = 4x^2+9$

259. $\int \frac{x}{\sqrt{4x^2+9}}dx = \frac{1}{4}\sqrt{4x^2+9} + C; u = 4x^2+9$

260. $\int \frac{x}{(4x^2+9)^2}dx = -\frac{1}{8(4x^2+9)}; u = 4x^2+9$

In the following exercises, find the antiderivative using the indicated substitution.

261. $\int (x+1)^4\,dx; u = x+1$

262. $\int (x-1)^5\,dx; u = x-1$

263. $\int (2x-3)^{-7}\,dx; u = 2x-3$

264. $\int (3x-2)^{-11}\,dx; u = 3x-2$

265. $\int \frac{x}{\sqrt{x^2+1}}dx; u = x^2+1$

266. $\int \frac{x}{\sqrt{1-x^2}}dx; u = 1-x^2$

267. $\int (x-1)\left(x^2-2x\right)^3 dx; u = x^2 - 2x$

268. $\int \left(x^2-2x\right)\left(x^3-3x^2\right)^2 dx; u = x^3 - 3x^2$

269. $\int \cos^3 \theta d\theta; u = \sin\theta \ (Hint: \cos^2\theta = 1 - \sin^2\theta)$

270. $\int \sin^3 \theta d\theta; u = \cos\theta \ (Hint: \sin^2\theta = 1 - \cos^2\theta)$

In the following exercises, use a suitable change of variables to determine the indefinite integral.

271. $\int x(1-x)^{99} dx$

272. $\int t\left(1-t^2\right)^{10} dt$

273. $\int (11x-7)^{-3} dx$

274. $\int (7x-11)^4 dx$

275. $\int \cos^3 \theta \sin\theta d\theta$

276. $\int \sin^7 \theta \cos\theta d\theta$

277. $\int \cos^2 (\pi t) \sin (\pi t) \, dt$

278. $\int \sin^2 x \cos^3 x dx \ (Hint: \sin^2 x + \cos^2 x = 1)$

279. $\int t \sin\left(t^2\right) \cos\left(t^2\right) dt$

280. $\int t^2 \cos^2\left(t^3\right) \sin\left(t^3\right) dt$

281. $\int \frac{x^2}{\left(x^3-3\right)^2}dx$

282. $\int \frac{x^3}{\sqrt{1-x^2}}dx$

283. $\int \frac{y^5}{\left(1-y^3\right)^{3/2}}dy$

284. $\int \cos\theta\,(1-\cos\theta)^{99}\sin\theta d\theta$

285. $\int \left(1-\cos^3\theta\right)^{10}\cos^2\theta\sin\theta d\theta$

286. $\int (\cos\theta-1)\left(\cos^2\theta-2\cos\theta\right)^3\sin\theta d\theta$

287. $\int \left(\sin^2\theta-2\sin\theta\right)\left(\sin^3\theta-3\sin^2\theta\right)^3\cos\theta d\theta$

In the following exercises, use a calculator to estimate the area under the curve using left Riemann sums with 50 terms, then use substitution to solve for the exact answer.

288. **[T]** $y=3(1-x)^2$ over $[0,2]$

289. **[T]** $y=x\left(1-x^2\right)^3$ over $[-1,2]$

290. **[T]** $y=\sin x(1-\cos x)^2$ over $[0,\pi]$

291. **[T]** $y=\frac{x}{\left(x^2+1\right)^2}$ over $[-1,1]$

In the following exercises, use a change of variables to evaluate the definite integral.

292. $\int_0^1 x\sqrt{1-x^2}dx$

293. $\int_0^1 \frac{x}{\sqrt{1+x^2}}dx$

294. $\int_0^2 \frac{t^2}{\sqrt{5+t^2}}dt$

295. $\int_0^1 \frac{t^2}{\sqrt{1+t^3}} dt$

296. $\int_0^{\pi/4} \sec^2\theta \tan\theta d\theta$

297. $\int_0^{\pi/4} \frac{\sin\theta}{\cos^4\theta} d\theta$

In the following exercises, evaluate the indefinite integral $\int f(x)\, dx$ with constant $C = 0$ using *u*-substitution. Then, graph the function and the antiderivative over the indicated interval. If possible, estimate a value of *C* that would need to be added to the antiderivative to make it equal to the definite integral $F(x) = \int_a^x f(t)\, dt$, with *a* the left endpoint of the given interval.

298. **[T]** $\int (2x+1) e^{x^2+x-6} dx$ over $[-3, 2]$

299. **[T]** $\int \frac{\cos(\ln(2x))}{x} dx$ on $[0, 2]$

300. **[T]** $\int \frac{3x^2+2x+1}{\sqrt{x^3+x^2+x+4}} dx$ over $[-1, 2]$

301. **[T]** $\int \frac{\sin x}{\cos^3 x} dx$ over $\left[-\frac{\pi}{3}, \frac{\pi}{3}\right]$

302. **[T]** $\int (x+2) e^{-x^2-4x+3} dx$ over $[-5, 1]$

303. **[T]** $\int 3x^2 \sqrt{2x^3+1} dx$ over $[0, 1]$

304. If $h(a) = h(b)$ in $\int_a^b g'(h(x))\, h(x)\, dx$, what can you say about the value of the integral?

305. Is the substitution $u = 1 - x^2$ in the definite integral $\int_0^2 \frac{x}{1-x^2} dx$ okay? If not, why not?

In the following exercises, use a change of variables to show that each definite integral is equal to zero.

306. $\int_0^{\pi} \cos^2(2\theta) \sin(2\theta)\, d\theta$

307. $\int_0^{\sqrt{\pi}} t \cos\left(t^2\right) \sin\left(t^2\right) dt$

308. $\int_0^1 (1 - 2t)\, dt$

309. $\int_0^1 \frac{1 - 2t}{\left(1 + \left(t - \frac{1}{2}\right)^2\right)} dt$

310. $\int_0^{\pi} \sin\left(\left(t - \frac{\pi}{2}\right)^3\right) \cos\left(t - \frac{\pi}{2}\right) dt$

311. $\int_0^2 (1 - t) \cos(\pi t)\, dt$

312. $\int_{\pi/4}^{3\pi/4} \sin^2 t \cos t\, dt$

313. Show that the average value of $f(x)$ over an interval $[a, b]$ is the same as the average value of $f(cx)$ over the interval $\left[\frac{a}{c}, \frac{b}{c}\right]$ for $c > 0$.

314. Find the area under the graph of $f(t) = \frac{t}{\left(1+t^2\right)^a}$ between $t = 0$ and $t = x$ where $a > 0$ and $a \neq 1$ is fixed, and evaluate the limit as $x \to \infty$.

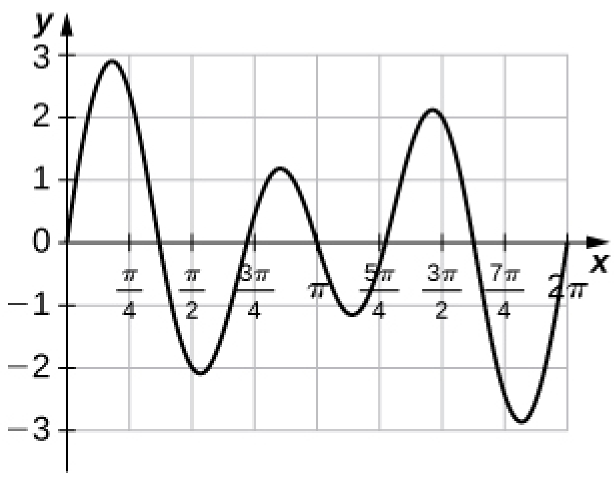
315. Find the area under the graph of $g(t) = \frac{t}{\left(1-t^2\right)^a}$ between $t = 0$ and $t = x$, where $0 < x < 1$ and $a > 0$ is fixed. Evaluate the limit as $x \to 1$.

316. The area of a semicircle of radius 1 can be expressed as $\int_{-1}^1 \sqrt{1 - x^2}\, dx.$ Use the substitution $x = \cos t$ to express the area of a semicircle as the integral of a trigonometric function. You do not need to compute the integral.

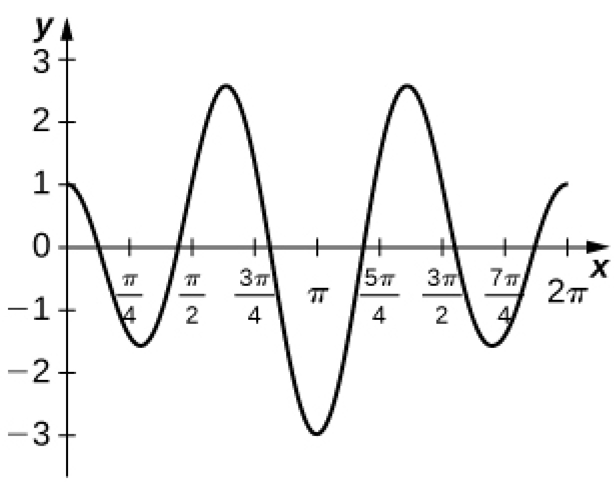
317. The area of the top half of an ellipse with a major axis that is the *x*-axis from $x = a$ to *a* and with a minor axis that is the *y*-axis from $y = -b$ to *b* can be written as $\int_{-a}^a b\sqrt{1 - \frac{x^2}{a^2}}\, dx$. Use the substitution $x = a \cos t$ to express this area in

terms of an integral of a trigonometric function. You do not need to compute the integral.

318. **[T]** The following graph is of a function of the form $f(t) = a\sin(nt) + b\sin(mt)$. Estimate the coefficients *a* and *b*, and the frequency parameters *n* and *m*. Use these estimates to approximate $\int_0^{\pi} f(t)\,dt$.




319. **[T]** The following graph is of a function of the form $f(x) = a\cos(nt) + b\cos(mt)$. Estimate the coefficients *a* and *b* and the frequency parameters *n* and *m*. Use these estimates to approximate $\int_0^{\pi} f(t)\,dt$.



## Learning Objectives

1.6.1. Integrate functions involving exponential functions.
1.6.2. Integrate functions involving logarithmic functions.

Exponential and logarithmic functions are used to model population growth, cell growth, and financial growth, as well as depreciation, radioactive decay, and resource consumption, to name only a few applications. In this section, we explore integration involving exponential and logarithmic functions.

## Integrals of Exponential Functions

The exponential function is perhaps the most efficient function in terms of the operations of calculus. The exponential function, $y = e^x$, is its own derivative and its own integral.

#### RULE: INTEGRALS OF EXPONENTIAL FUNCTIONS

Exponential functions can be integrated using the following formulas.

$$\begin{aligned} \int e^x\,dx &= e^x + C \\ \int a^x\,dx &= \frac{a^x}{\ln a} + C \end{aligned} \tag{1.21}$$

#### EXAMPLE 1.37

### Finding an Antiderivative of an Exponential Function

Find the antiderivative of the exponential function $e^{-x}$.

[Show Solution]

#### CHECKPOINT 1.31

Find the antiderivative of the function using substitution: $x^2 e^{-2x^3}$.

A common mistake when dealing with exponential expressions is treating the exponent on *e* the same way we treat exponents in polynomial expressions. We cannot use the power rule for the exponent on *e*. This can be especially confusing when we have both exponentials and polynomials in the same expression, as in the previous checkpoint. In these cases, we should always double-check to make sure we're using the right rules for the functions we're integrating.

**EXAMPLE 1.38**

### Square Root of an Exponential Function

Find the antiderivative of the exponential function $e^x \sqrt{1 + e^x}$.

[Show Solution]

**CHECKPOINT 1.32**

Find the antiderivative of $e^x (3e^x - 2)^2$.

**EXAMPLE 1.39**

### Using Substitution with an Exponential Function

Use substitution to evaluate the indefinite integral $\int 3x^2 e^{2x^3} dx$.

[Show Solution]

### CHECKPOINT 1.33

Evaluate the indefinite integral $\int 2x^3 e^{x^4} dx$.

As mentioned at the beginning of this section, exponential functions are used in many real-life applications. The number *e* is often associated with compounded or accelerating growth, as we have seen in earlier sections about the derivative. Although the derivative represents a rate of change or a growth rate, the integral represents the total change or the total growth. Let's look at an example in which integration of an exponential function solves a common business application.

A price–demand function tells us the relationship between the quantity of a product demanded and the price of the product. In general, price decreases as quantity demanded increases. The marginal price–demand function is the derivative of the price–demand function and it tells us how fast the price changes at a given level of production. These functions are used in business to determine the price–elasticity of demand, and to help companies determine whether changing production levels would be profitable.

### EXAMPLE 1.40

#### Finding a Price–Demand Equation

Find the price–demand equation for a particular brand of toothpaste at a supermarket chain when the demand is 50 tubes per week at \$2.35 per tube, given that the marginal price—demand function, $p'(x)$, for *x* number of tubes per week, is given as

$$p'(x) = -0.015e^{-0.01x}.$$

If the supermarket chain sells 100 tubes per week, what price should it set?

[Show Solution]

EXAMPLE 1.41

### Evaluating a Definite Integral Involving an Exponential Function

Evaluate the definite integral $\int_1^2 e^{1-x}\,dx$.

[Show Solution]

CHECKPOINT 1.34

Evaluate $\int_0^2 e^{2x}\,dx$.

EXAMPLE 1.42

### Growth of Bacteria in a Culture

Suppose the rate of growth of bacteria in a Petri dish is given by $q(t) = 3^t$, where $t$ is given in hours and $q(t)$ is given in thousands of bacteria per hour. If a culture starts with 10,000 bacteria, find a function $Q(t)$ that gives the number of bacteria in the Petri dish at any time $t$. How many bacteria are in the dish after 2 hours?

[Show Solution]

CHECKPOINT 1.35

From Example 1.42, suppose the bacteria grow at a rate of $q(t) = 2^t$. Assume the culture still starts with 10,000 bacteria. Find $Q(t)$. How many bacteria are in the dish after 3 hours?

EXAMPLE 1.43

### Fruit Fly Population Growth

Suppose a population of fruit flies increases at a rate of $g(t) = 2e^{0.02t}$, in flies per day. If the initial population of fruit flies is 100 flies, how many flies are in the population after 10 days?

[Show Solution]

CHECKPOINT 1.36

Suppose the rate of growth of the fly population is given by $g(t) = e^{0.01t}$, and the initial fly population is 100 flies. How many flies are in the population after 15 days?

EXAMPLE 1.44

### Evaluating a Definite Integral Using Substitution

Evaluate the definite integral using substitution: $\int_1^2 \frac{e^{1/x}}{x^2}\,dx$.

[Show Solution]

### CHECKPOINT 1.37

Evaluate the definite integral using substitution: $\int_1^2 \frac{1}{x^3} e^{4x^{-2}} dx.$

## Integrals Involving Logarithmic Functions

Integrating functions of the form $f(x) = x^{-1}$ result in the absolute value of the natural log function, as shown in the following rule. Integral formulas for other logarithmic functions, such as $f(x) = \ln x$ and $f(x) = \log_a x$, are also included in the rule.

### RULE: INTEGRATION FORMULAS INVOLVING LOGARITHMIC FUNCTIONS

The following formulas can be used to evaluate integrals involving logarithmic functions.

$$\begin{aligned} \int x^{-1} dx &= \ln|x| + C \\ \int \ln x \, dx &= x \ln x - x + C = x(\ln x - 1) + C \\ \int \log_a x \, dx &= \frac{x}{\ln a}(\ln x - 1) + C \end{aligned} \tag{1.22}$$

### EXAMPLE 1.45

#### Finding an Antiderivative Involving $\ln x$

Find the antiderivative of the function $\frac{3}{x-10}$.

[Show Solution]

### CHECKPOINT 1.38

Find the antiderivative of $\frac{1}{x+2}$.

### EXAMPLE 1.46

#### Finding an Antiderivative of a Rational Function

Find the antiderivative of $\frac{2x^3+3x}{x^4+3x^2}$.

[Show Solution]

### EXAMPLE 1.47

#### Finding an Antiderivative of a Logarithmic Function

Find the antiderivative of the log function $\log_2 x$.

[Show Solution]

### CHECKPOINT 1.39

Find the antiderivative of $\log_3 x$.

Example 1.48 is a definite integral of a trigonometric function. With trigonometric functions, we often have to apply a trigonometric property or an identity before we can move forward. Finding the right form of the integrand is usually the key to a smooth integration.

**EXAMPLE 1.48**

### Evaluating a Definite Integral

Find the definite integral of $\int_0^{\pi/2} \frac{\sin x}{1 + \cos x} dx.$

[Show Solution]

## Section 1.6 Exercises

In the following exercises, compute each indefinite integral.

320. $\int e^{2x}\, dx$

321. $\int e^{-3x}\, dx$

322. $\int 2^x\, dx$

323. $\int 3^{-x}\, dx$

324. $\int \frac{1}{2x} dx$

325. $\int \frac{2}{x} dx$

326. $\int \frac{1}{x^2} dx$

327. $\int \frac{1}{\sqrt{x}} dx$

In the following exercises, find each indefinite integral by using appropriate substitutions.

328. $\int \frac{\ln x}{x} dx$

329. $\displaystyle\int \frac{dx}{x(\ln x)^2}$

330. $\displaystyle\int \frac{dx}{x \ln x}\ (x > 1)$

331. $\displaystyle\int \frac{dx}{x \ln x \ln(\ln x)}$

332. $\displaystyle\int \tan\theta\ d\theta$

333. $\displaystyle\int \frac{\cos x - x \sin x}{x \cos x} dx$

334. $\displaystyle\int \frac{\ln(\sin x)}{\tan x} dx$

335. $\displaystyle\int \ln(\cos x) \tan x dx$

336. $\displaystyle\int x e^{-x^2} dx$

337. $\displaystyle\int x^2 e^{-x^3} dx$

338. $\displaystyle\int e^{\sin x} \cos x dx$

339. $\displaystyle\int e^{\tan x} \sec^2 x dx$

340. $\displaystyle\int e^{\ln x} \frac{dx}{x}$

341. $\displaystyle\int \frac{e^{\ln(1-t)}}{1-t} dt$

In the following exercises, verify by differentiation that $\displaystyle\int \ln x\ dx = x(\ln x - 1) + C$, then use appropriate changes of variables to compute the integral.

342. $\displaystyle\int \ln x dx\ (Hint\text{:}\ \int \ln x dx = \frac{1}{2}\int x \ln\left(x^2\right) dx)$

343. $\int x^2 \ln^2 x \, dx$

344. $\int \frac{\ln x}{x^2} dx$ ($Hint$: Set $u = \frac{1}{x}$.)

345. $\int \frac{\ln x}{\sqrt{x}} dx$ ($Hint$: Set $u = \sqrt{x}$.)

346. Write an integral to express the area under the graph of $y = \frac{1}{t}$ from $t = 1$ to $e^x$ and evaluate the integral.

347. Write an integral to express the area under the graph of $y = e^t$ between $t = 0$ and $t = \ln x$, and evaluate the integral.

In the following exercises, use appropriate substitutions to express the trigonometric integrals in terms of compositions with logarithms.

348. $\int \tan(2x) \, dx$

349. $\int \frac{\sin(3x) - \cos(3x)}{\sin(3x) + \cos(3x)} dx$

350. $\int \frac{x \sin\left(x^2\right)}{\cos\left(x^2\right)} dx$

351. $\int x \csc\left(x^2\right) dx$

352. $\int \ln(\cos x) \tan x \, dx$

353. $\int \ln(\csc x) \cot x dx$

354. $\int \frac{e^x - e^{-x}}{e^x + e^{-x}} dx$

In the following exercises, evaluate the definite integral.

355. $\int_1^2 \frac{1 + 2x + x^2}{3x + 3x^2 + x^3} dx$

356. $\int_0^{\pi/4} \tan x \, dx$

357. $\int_0^{\pi/3} \frac{\sin x - \cos x}{\sin x + \cos x} dx$

358. $\int_{\pi/6}^{\pi/2} \csc x dx$

359. $\int_{\pi/4}^{\pi/3} \cot x dx$

In the following exercises, integrate using the indicated substitution.

360. $\int \frac{x}{x-100} dx; u = x - 100$

361. $\int \frac{y-1}{y+1} dy; u = y + 1$

362. $\int \frac{1-x^2}{3x-x^3} dx; u = 3x - x^3$

363. $\int \frac{\sin x + \cos x}{\sin x - \cos x} dx; u = \sin x - \cos x$

364. $\int e^{2x} \sqrt{1-e^{2x}} dx; u = e^{2x}$

365. $\int \ln(x) \frac{\sqrt{1-(\ln x)^2}}{x} dx; u = \ln x$

In the following exercises, does the right-endpoint approximation overestimate or underestimate the exact area? Calculate the right endpoint estimate $R_{50}$ and solve for the exact area.

366. **[T]** $y = e^x$ over $[0, 1]$

367. **[T]** $y = e^{-x}$ over $[0, 1]$

368. **[T]** $y = \ln(x)$ over $[1, 2]$

369. **[T]** $y = \frac{x+1}{x^2+2x+6}$ over $[0, 1]$

370. **[T]** $y = 2^x$ over $[-1, 0]$

371. **[T]** $y = -2^{-x}$ over $[0, 1]$

In the following exercises, $f(x) \geq 0$ for $a \leq x \leq b$. Find the area under the graph of $f(x)$ between the given values *a* and *b* by integrating.

372. $f(x) = \frac{\log_{10}(x)}{x}; a = 10, b = 100$

373. $f(x) = \frac{\log_2(x)}{x}; a = 32, b = 64$

374. $f(x) = 2^{-x}; a = 1, b = 2$

375. $f(x) = 2^{-x}; a = 3, b = 4$

376. Find the area under the graph of the function $f(x) = xe^{-x^2}$ between $x = 0$ and $x = 5$.

377. Compute the integral of $f(x) = xe^{-x^2}$ and find the smallest value of *N* such that the area under the graph $f(x) = xe^{-x^2}$ between $x = N$ and $x = N + 1$ is, at most, 0.01.

378. Find the limit, as *N* tends to infinity, of the area under the graph of $f(x) = xe^{-x^2}$ between $x = 0$ and $x = 5$.

379. Show that $\int_a^b \frac{dt}{t} = \int_{1/b}^{1/a} \frac{dt}{t}$ when $0 < a \leq b$.

380. Suppose that $f(x) > 0$ for all *x* and that *f* and *g* are differentiable. Use the identity $f^g = e^{g \ln f}$ and the chain rule to find the derivative of $f^g$.

381. Use the previous exercise to find the antiderivative of $h(x) = x^x(1 + \ln x)$ and evaluate $\int_2^3 x^x(1 + \ln x)\,dx$.

382. Show that if $c > 0$, then the integral of $1/x$ from *ac* to *bc* $(0 < a < b)$ is the same as the integral of $1/x$ from *a* to *b*.

The following exercises are intended to derive the fundamental properties of the natural log starting from the *definition* $\ln(x) = \int_1^x \frac{dt}{t}$, using properties of the definite integral and making no further assumptions.

383. Use the identity $\ln(x) = \int_1^x \frac{dt}{t}$ to derive the identity $\ln\left(\frac{1}{x}\right) = -\ln x$.

384. Use a change of variable in the integral $\int_1^{xy} \frac{1}{t}\,dt$ to show that $\ln xy = \ln x + \ln y$ for $x, y > 0$.

385. Use the identity $\ln x = \int_1^x \frac{dt}{x}$ to show that $\ln(x)$ is an increasing function of *x* on $[0, \infty)$, and use the previous exercises to show that the range of $\ln(x)$ is $(-\infty, \infty)$. Without any further assumptions, conclude that $\ln(x)$ has an inverse function defined on $(-\infty, \infty)$.

386. Pretend, for the moment, that we do not know that $e^x$ is the inverse function of $\ln(x)$, but keep in mind that $\ln(x)$ has an inverse function defined on $(-\infty, \infty)$. Call it *E*. Use the identity $\ln xy = \ln x + \ln y$ to deduce that $E(a+b) = E(a)E(b)$ for any real numbers *a*, *b*.

387. Pretend, for the moment, that we do not know that $e^x$ is the inverse function of $\ln x$, but keep in mind that $\ln x$ has an inverse function defined on $(-\infty, \infty)$. Call it *E*. Show that $E'(t) = E(t)$.

388. The sine integral, defined as $S(x) = \int_0^x \frac{\sin t}{t} dt$ is an important quantity in engineering. Although it does not have a simple closed formula, it is possible to estimate its behavior for large *x*. Show that for $k \geq 1, |S(2\pi k) - S(2\pi(k+1))| \leq \frac{1}{k(2k+1)\pi}$. $(Hint\text{: } \sin(t+\pi) = -\sin t)$

389. **[T]** The normal distribution in probability is given by $p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$, where $\sigma$ is the standard deviation and $\mu$ is the average. The *standard normal distribution* in probability, $p_s$, corresponds to $\mu = 0$ and $\sigma = 1$. Compute the right endpoint estimates $R_{10}$ and $R_{100}$ of $\int_{-1}^1 \frac{1}{\sqrt{2\pi}} e^{-x^{2/2}} dx$.

390. **[T]** Compute the right endpoint estimates $R_{50}$ and $R_{100}$ of $\int_{-3}^5 \frac{1}{2\sqrt{2\pi}} e^{-(x-1)^2/8}$.

## Learning Objectives

3.1.1. Recognize when to use integration by parts.
3.1.2. Use the integration-by-parts formula to solve integration problems.
3.1.3. Use the integration-by-parts formula for definite integrals.

By now we have a fairly thorough procedure for how to evaluate many basic integrals. However, although we can integrate $\int x\sin(x^2)dx$ by using the substitution, $u = x^2$, something as simple looking as $\int x\sin x\,dx$ defies us. Many students want to know whether there is a product rule for integration. There isn't, but there is a technique based on the product rule for differentiation that allows us to exchange one integral for another. We call this technique **integration by parts**.

## The Integration-by-Parts Formula

If, $h(x) = f(x)g(x)$, then by using the product rule, we obtain $h'(x) = f'(x)g(x) + g'(x)f(x)$. Although at first it may seem counterproductive, let's now integrate both sides of this equation:
$\int h'(x)dx = \int \left(g(x)f'(x) + f(x)g'(x)\right)dx.$

This gives us

$$h(x) = f(x)g(x) = \int g(x)f'(x)dx + \int f(x)g'(x)dx.$$

Now we solve for $\int f(x)g'(x)dx$ :

$$\int f(x)g'(x)dx = f(x)g(x) - \int g(x)f'(x)dx.$$

By making the substitutions $u = f(x)$ and $v = g(x)$, which in turn make $du = f'(x)dx$ and $dv = g'(x)dx$, we have the more compact form

$$\int u\,dv = uv - \int v\,du.$$

**THEOREM 3.1**

### Integration by Parts

Let $u = f(x)$ and $v = g(x)$ be functions with continuous derivatives. Then, the integration-by-parts formula for the integral involving these two functions is:

$$\int u\,dv = uv - \int v\,du. \tag{3.1}$$

The advantage of using the integration-by-parts formula is that we can use it to exchange one integral for another, possibly easier, integral. The following example illustrates its use.

**EXAMPLE 3.1**

### Using Integration by Parts

Use integration by parts with $u = x$ and $dv = \sin x\,dx$ to evaluate $\int x \sin x\,dx.$

[Show Solution]

### Analysis

At this point, there are probably a few items that need clarification. First of all, you may be curious about what would have happened if we had chosen $u = \sin x$ and $dv = x$. If we had done so, then we would have $du = \cos x\,\mathrm{dx}$ and $v = \frac{1}{2}x^2$. Thus, after applying integration by parts, we have $\int x \sin x\,dx = \frac{1}{2}x^2 \sin x - \int \frac{1}{2}x^2 \cos x\,dx.$ Unfortunately, with the new integral, we are in no better position than before. It is important to keep in mind that when we apply integration by parts, we may need to try several choices for $u$ and $dv$ before finding a choice that works.

Second, you may wonder why, when we find $v = \int \sin x\,dx = -\cos x$, we do not use $v = -\cos x + K$. To see that it makes no difference, we can rework the problem using $v = -\cos x + K$:

$$\begin{aligned}\int x\sin x\,dx &= (x)(-\cos x + K) - \int (-\cos x + K)(1dx) \\ &= -x\cos x + Kx + \int \cos x\,dx - \int Kdx \\ &= -x\cos x\ + Kx + \sin x - Kx + C \\ &= -x\cos x + \sin x + C.\end{aligned}$$

As you can see, it makes no difference in the final solution.

Last, we can check to make sure that our antiderivative is correct by differentiating $-x\cos x + \sin x + C$:

$$\begin{aligned}\tfrac{d}{dx}(-x\cos x + \sin x + C) &= (-1)\cos x + (-x)(-\sin x) + \cos x \\ &= x\sin x.\end{aligned}$$

Therefore, the antiderivative checks out.

## MEDIA

Watch this video and visit this website for examples of integration by parts.

## CHECKPOINT 3.1

Evaluate $\int xe^{2x}\,dx$ using the integration-by-parts formula with $u = x$ and $dv = e^{2x}\,dx$.

The natural question to ask at this point is: How do we know how to choose $u$ and $dv$? Sometimes it is a matter of trial and error; however, the acronym LIATE can often help to take some of the guesswork out of our choices. This acronym stands for **L**ogarithmic Functions, **I**nverse Trigonometric Functions, **A**lgebraic Functions, **T**rigonometric Functions, and **E**xponential Functions. This mnemonic serves as an aid in determining an appropriate choice for $u$.

The type of function in the integral that appears first in the list should be our first choice of $u$. For example, if an integral contains a logarithmic function and an algebraic function, we should choose $u$ to be the logarithmic function, because L comes before A in LIATE. The integral in Example 3.1 has a trigonometric function $(\sin x)$ and an algebraic function $(x)$. Because A comes before T in LIATE, we chose $u$ to be the algebraic function. When we have chosen $u$, $dv$ is selected to be the remaining part of the function to be integrated, together with $dx$.

Why does this mnemonic work? Remember that whatever we pick to be $dv$ must be something we can integrate. Since we do not have integration formulas that allow us to integrate simple logarithmic functions and inverse trigonometric functions, it makes sense that they should not be chosen as values for $dv$. Consequently, they should be at the head of the list as choices for $u$. Thus, we put LI at the beginning of the mnemonic. (We could just as easily have started with IL, since these two types of functions won't appear together in an integration-by-parts problem.) The exponential and trigonometric functions are at the end of our list because they are fairly easy to integrate and make good choices for $dv$. Thus, we have TE at the end of our mnemonic. (We could just as easily have used ET at the end, since when these types of functions appear together it usually doesn't really matter which one is $u$ and which one is $dv$.) Algebraic functions are generally easy both to integrate and to differentiate, and they come in the middle of the mnemonic.

**EXAMPLE 3.2**

**Using Integration by Parts**

Evaluate $\int \frac{\ln x}{x^3} dx$.

**[Show Solution]**

**CHECKPOINT 3.2**

Evaluate $\int x \ln x \, dx$.

In some cases, as in the next two examples, it may be necessary to apply integration by parts more than once.

### EXAMPLE 3.3

#### Applying Integration by Parts More Than Once

Evaluate $\int x^2 e^{3x}\,dx.$

[Show Solution]

### EXAMPLE 3.4

#### Applying Integration by Parts When LIATE Doesn't Quite Work

Evaluate $\int t^3 e^{t^2}\,dt.$

[Show Solution]

### EXAMPLE 3.5

#### Applying Integration by Parts More Than Once

Evaluate $\int \sin(\ln x)\,dx.$

[Show Solution]

#### Analysis

If this method feels a little strange at first, we can check the answer by differentiation:

$$\frac{d}{dx}\left(\frac{1}{2}x\sin(\ln x) - \frac{1}{2}x\cos(\ln x)\right)$$

$$= \frac{1}{2}(\sin(\ln x)) + \cos(\ln x) \cdot \frac{1}{x} \cdot \frac{1}{2}x - \left(\frac{1}{2}\cos(\ln x) - \sin(\ln x) \cdot \frac{1}{x} \cdot \frac{1}{2}x\right)$$
$$= \sin(\ln x).$$

**CHECKPOINT 3.3**

Evaluate $\displaystyle\int x^2 \sin x \, dx.$

## Integration by Parts for Definite Integrals

Now that we have used integration by parts successfully to evaluate indefinite integrals, we turn our attention to definite integrals. The integration technique is really the same, only we add a step to evaluate the integral at the upper and lower limits of integration.

**THEOREM 3.2**

### Integration by Parts for Definite Integrals

Let $u = f(x)$ and $v = g(x)$ be functions with continuous derivatives on $[a, b]$. Then

$$\int_a^b u \, dv = uv\Big|_a^b - \int_a^b v \, du. \tag{3.2}$$

**EXAMPLE 3.6**

### Finding the Area of a Region

Find the area of the region bounded above by the graph of $y = \tan^{-1} x$ and below by the $x$-axis over the interval $[0, 1]$.

[Show Solution]

EXAMPLE 3.7

### Finding a Volume of Revolution

Find the volume of the solid obtained by revolving the region bounded by the graph of $f(x) = e^{-x}$, the *x*-axis, the *y*-axis, and the line $x = 1$ about the *y*-axis.

[Show Solution]

### Analysis

Again, it is a good idea to check the reasonableness of our solution. We observe that the solid has a volume slightly less than that of a cylinder of radius $1$ and height of $1/e$ added to the volume of a cone of base radius $1$ and height of $1 - \frac{1}{e}$. Consequently, the solid should have a volume a bit less than

$$\pi(1)^2 \frac{1}{e} + \left(\frac{\pi}{3}\right)(1)^2 \left(1 - \frac{1}{e}\right) = \frac{2\pi}{3e} + \frac{\pi}{3} \approx 1.8177.$$

Since $2\pi - \frac{4\pi}{e} \approx 1.6603$, we see that our calculated volume is reasonable.

CHECKPOINT 3.4

Evaluate $\displaystyle\int_0^{\pi/2} x \cos x \, dx.$

## Section 3.1 Exercises

In using the technique of integration by parts, you must carefully choose which expression is *u.* For each of the following problems, use the guidelines in this section to choose *u.* Do **not** evaluate the integrals.

1. $\int x^3 e^{2x}\,dx$
2. $\int x^3 \ln(x) dx$
3. $\int y^3 \cos y dy$
4. $\int x^2 \arctan x\ dx$
5. $\int e^{3x} \sin(2x) dx$

Find the integral by using the simplest method. Not all problems require integration by parts.

6. $\int v \sin v dv$
7. $\int \ln x\ dx$ (*Hint:* $\int \ln x\ dx$ is equivalent to $\int 1 \cdot \ln(x) dx.$)
8. $\int x \cos x\ dx$
9. $\int \tan^{-1} x\ dx$
10. $\int x^2 e^x\,dx$
11. $\int x \sin(2x) dx$
12. $\int x e^{4x}\,dx$
13. $\int x e^{-x}\,dx$
14. $\int x \cos 3x\ dx$

15. $\int x^2 \cos x \, dx$

16. $\int x \ln x \, dx$

17. $\int \ln(2x + 1) dx$

18. $\int x^2 e^{4x} dx$

19. $\int e^x \sin x \, dx$

20. $\int e^x \cos x \, dx$

21. $\int x e^{-x^2} dx$

22. $\int x^2 e^{-x} dx$

23. $\int \sin(\ln(2x)) dx$

24. $\int cos(\ln x) dx$

25. $\int (\ln x)^2 dx$

26. $\int \ln(x^2) dx$

27. $\int x^2 \ln x \, dx$

28. $\int \sin^{-1} x \, dx$

29. $\int \cos^{-1}(2x) dx$

30. $\int x \arctan x \, dx$

31. $\int x^2 \sin x \, dx$

32. $\int x^3 \cos x \, dx$

33. $\int x^3 \sin x \, dx$

34. $\int x^3 e^x \, dx$

35. $\int x \sec^{-1} x \, dx$

36. $\int x \sec^2 x \, dx$

37. $\int x \cosh x \, dx$

Compute the definite integrals. Use a graphing utility to confirm your answers.

38. $\int_{1/e}^{1} \ln x \, dx$

39. $\int_0^1 xe^{-2x} dx$ (Express the answer in exact form.)

40. $\int_0^1 e^{\sqrt{x}} dx \left(\text{let } u = \sqrt{x}\right)$

41. $\int_1^e \ln(x^2) dx$

42. $\int_0^{\pi} x \cos x \, dx$

43. $\int_{-\pi}^{\pi} x \sin x \, dx$ (Express the answer in exact form.)

44. $\int_0^3 \ln(x^2 + 1) dx$ (Express the answer in exact form.)

45. $\int_0^{\pi/2} x^2 \sin x \, dx$ (Express the answer in exact form.)

46. $\int_0^1 x5^x\,dx$ (Express the answer using five significant digits.)

47. Evaluate $\int \cos x \ln(\sin x) dx$

Derive the following formulas using the technique of integration by parts. Assume that *n* is a positive integer. These formulas are called *reduction formulas* because the exponent in the *x* term has been reduced by one in each case. The second integral is simpler than the original integral.

48. $\int x^n e^x\,dx = x^n e^x - n \int x^{n-1} e^x\,dx$

49. $\int x^n \cos x\ dx = x^n \sin x - n \int x^{n-1} \sin x\ dx$

50. $\int x^n \sin x\ dx =$______

51. Integrate $\int 2x\sqrt{2x-3}dx$ using two methods:

a. Using parts, letting $dv = \sqrt{2x-3}\,dx$
b. Substitution, letting $u = 2x - 3$

State whether you would use integration by parts to evaluate the integral. If so, identify *u* and *dv*. If not, describe the technique used to perform the integration without actually doing the problem.

52. $\int x \ln x\ dx$

53. $\int \frac{\ln^2 x}{x}\,dx$

54. $\int xe^x\,dx$

55. $\int xe^{x^2-3}\,dx$

56. $\int x^2 \sin x\ dx$

57. $\int x^2 \sin(3x^3 + 2)dx$

Sketch the region bounded above by the curve, the *x*-axis, and $x = 1$, and find the area of the region. Provide the exact form or round answers to the number of places indicated.

58. $y = 2xe^{-x}$ (Approximate answer to four decimal places.)

59. $y = e^{-x}\sin(\pi x)$ (Approximate answer to five decimal places.)

Find the volume generated by rotating the region bounded by the given curves about the specified line. Express the answers in exact form or approximate to the number of decimal places indicated.

60. $y = \sin x, y = 0, x = 2\pi, x = 3\pi$ about the y-axis (Express the answer in exact form.)

61. $y = e^{-x}$ $y = 0, x = -1x = 0$; about $x = 1$ (Express the answer in exact form.)

62. A particle moving along a straight line has a velocity of $v(t) = t^2e^{-t}$ after *t* sec. How far does it travel in the first 2 sec? (Assume the units are in feet and express the answer in exact form.)

63. Find the area under the graph of $y = \sec^3 x$ from $x = 0$ to $x = 1$. (Round the answer to two significant digits.)

64. Find the area between $y = (x - 2)e^x$ and the *x*-axis from $x = 2$ to $x = 5$. (Express the answer in exact form.)

65. Find the area of the region enclosed by the curve $y = x\cos x$ and the *x*-axis for $\frac{11\pi}{2} \leq x \leq \frac{13\pi}{2}$. (Express the answer in exact form.)

66. Find the volume of the solid generated by revolving the region bounded by the curve $y = \ln x$, the *x*-axis, and the vertical line $x = e^2$ about the *x*-axis. (Express the answer in exact form.)

67. Find the volume of the solid generated by revolving the region bounded by the curve $y = 4\cos x$ and the *x*-axis, $\frac{\pi}{2} \leq x \leq \frac{3\pi}{2}$, about the *x*-axis. (Express the answer in exact form.)

68. Find the volume of the solid generated by revolving the region in the first quadrant bounded by $y = e^x$ and the *x*-axis, from $x = 0$ to $x = \ln(7)$, about the *y*-axis. (Express the answer in exact form.)

**Figure 5.1** The heights of these radish plants are continuous random variables. (Credit: Rev Stan)

### CHAPTER OBJECTIVES

By the end of this chapter, the student should be able to:

- Recognize and understand continuous probability density functions in general.
- Recognize the uniform probability distribution and apply it appropriately.
- Recognize the exponential probability distribution and apply it appropriately.

Continuous random variables have many applications. Baseball batting averages, IQ scores, the length of time a long distance telephone call lasts, the amount of money a person carries, the length of time a computer chip lasts, and SAT scores are just a few. The field of reliability depends on a variety of continuous random variables.

> **NOTE**
>
> The values of discrete and continuous random variables can be ambiguous. For example, if *X* is equal to the number of miles (to the nearest mile) you drive to work, then *X* is a discrete random variable. You count the miles. If *X* is the distance you drive to work, then you measure values of *X* and *X* is a continuous random variable. For a second example, if *X* is equal to the number of books in a backpack, then *X* is a discrete random variable. If *X* is the weight of a book, then *X* is a continuous random variable because weights are measured. How the random variable is defined is very important.

## Properties of Continuous Probability Distributions

The graph of a continuous probability distribution is a curve. Probability is represented by area under the curve.

The curve is called the **probability density function** (abbreviated as **pdf**). We use the symbol $f(x)$ to represent the curve. $f(x)$ is the function that corresponds to the graph; we use the density function $f(x)$ to draw the graph of the probability distribution.

**Area under the curve** is given by a different function called the **cumulative distribution function** (abbreviated as **cdf**). The cumulative distribution function is used to evaluate probability as area.

- The outcomes are measured, not counted.
- The entire area under the curve and above the x-axis is equal to one.
- Probability is found for intervals of *x* values rather than for individual *x* values.
- $P(c < x < d)$ is the probability that the random variable *X* is in the interval between the values *c* and *d*. $P(c < x < d)$ is the area under the curve, above the *x*-axis, to the right of *c* and the left of *d*.
- $P(x = c) = 0$ The probability that *x* takes on any single individual value is zero. The area below the curve, above the *x*-axis, and between $x = c$ and $x = c$ has no width, and therefore no area (area = 0). Since the probability is equal to the area, the probability is also zero.
- $P(c < x < d)$ is the same as $P(c \leq x \leq d)$ because probability is equal to area.

We will find the area that represents probability by using geometry, formulas, technology, or probability tables. In general, calculus is needed to find the area under the curve for many probability density functions. When we use formulas to find the area in this textbook, the formulas were found by using the techniques of integral calculus. However, because most students taking this course have not studied calculus, we will not be using calculus in this textbook.

There are many continuous probability distributions. When using a continuous probability distribution to model probability, the distribution used is selected to model and fit the particular situation in the best way.

In this chapter and the next, we will study the uniform distribution, the exponential distribution, and the normal distribution. The following graphs illustrate these distributions.

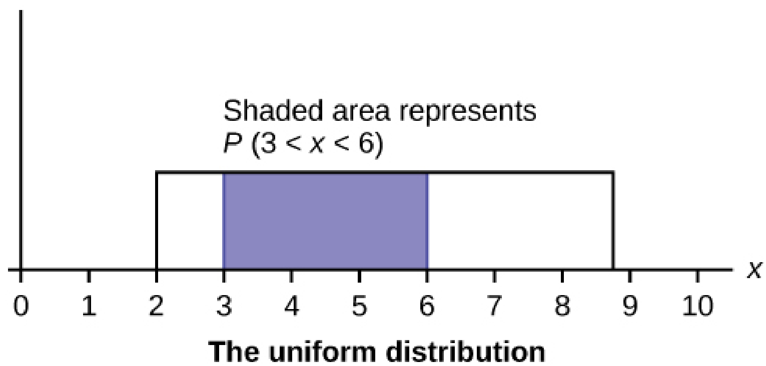



**Figure 5.2** The graph shows a Uniform Distribution with the area between $x = 3$ and $x = 6$ shaded to represent the probability that the value of the random variable $X$ is in the interval between three and six.

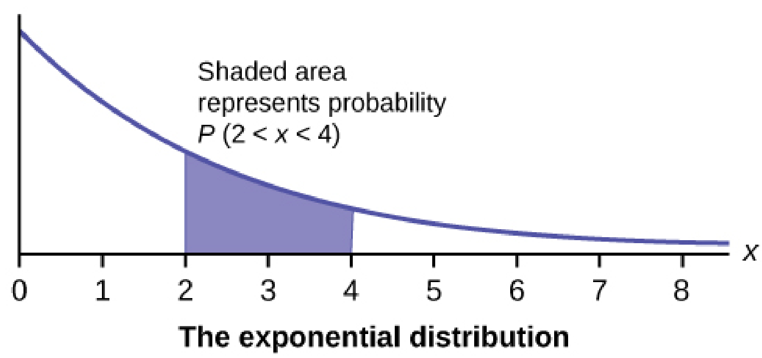



**Figure 5.3** The graph shows an Exponential Distribution with the area between $x = 2$ and $x = 4$ shaded to represent the probability that the value of the random variable $X$ is in the interval between two and four.

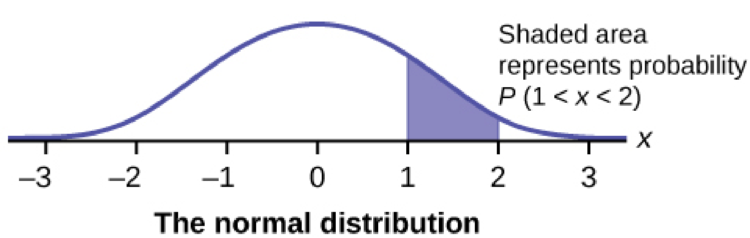



**Figure 5.4** The graph shows the Standard Normal Distribution with the

area between $x = 1$ and $x = 2$ shaded to represent the probability that the value of the random variable $X$ is in the interval between one and two.

We begin by defining a continuous probability density function. We use the function notation *f*(*x*). Intermediate algebra may have been your first formal introduction to functions. In the study of probability, the functions we study are special. We define the function *f*(*x*) so that the area between it and the x-axis is equal to a probability. Since the maximum probability is one, the maximum area is also one. **For continuous probability distributions, PROBABILITY = AREA.**

### EXAMPLE 5.1

Consider the function $f(x) = \frac{1}{20}$ for $0 \le x \le 20$. $x$ = a real number. The graph of $f(x) = \frac{1}{20}$ is a horizontal line. However, since $0 \le x \le 20$, $f(x)$ is restricted to the portion between $x = 0$ and $x = 20$, inclusive.

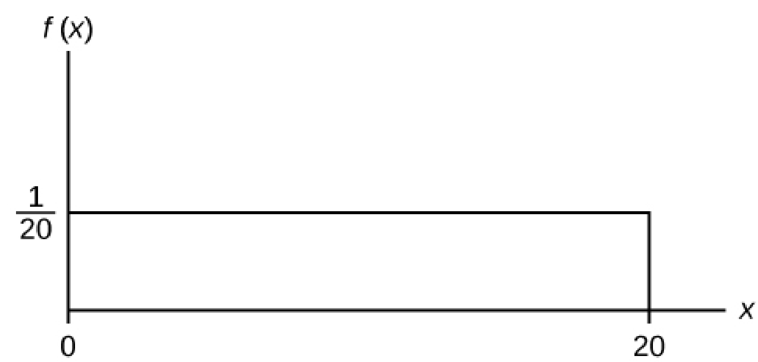



**Figure 5.5**

$f(x) = \frac{1}{20}$ **for** $0 \le x \le 20$.

The graph of $f(x) = \frac{1}{20}$ is a horizontal line segment when $0 \le x \le 20$.

The area between $f(x) = \frac{1}{20}$ where $0 \le x \le 20$ and the $x$-axis is the area of a rectangle with base = 20 and height = $\frac{1}{20}$.

$$\text{AREA} = 20\left(\frac{1}{20}\right) = 1$$

**Suppose we want to find the area between $f(x) = \frac{1}{20}$ and the $x$-axis where $0 < x < 2$.**

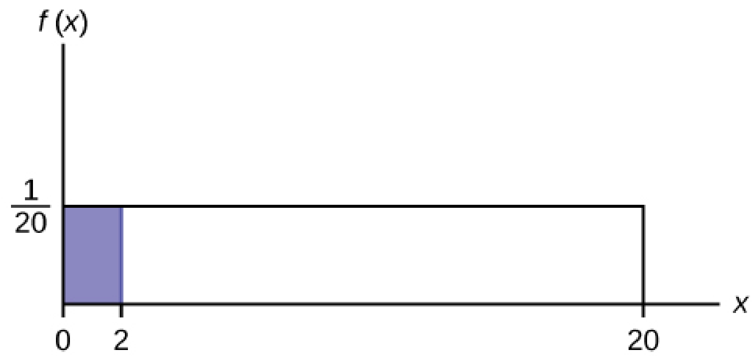



Figure 5.6

$$\text{AREA} = (2-0)\left(\frac{1}{20}\right) = 0.1$$

$$(2-0) = 2 = \text{base of a rectangle}$$

**REMINDER**

area of a rectangle = (base)(height).

The area corresponds to a probability. The probability that $x$ is between zero and two is 0.1, which can be written mathematically as $P(0 < x < 2) = P(x < 2) = 0.1$.

**Suppose we want to find the area between $f(x) = \frac{1}{20}$ and the $x$-axis where $4 < x < 15$.**

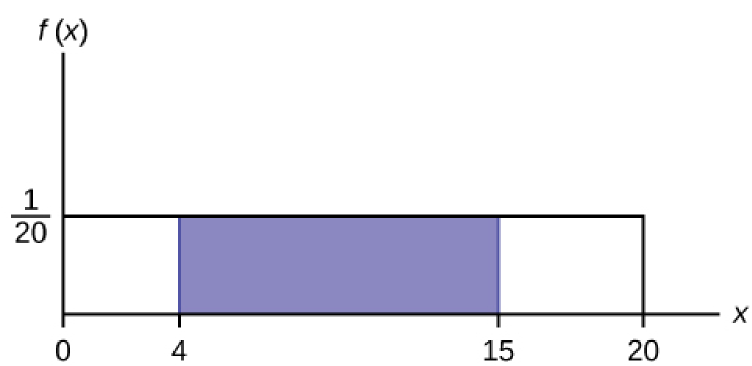



Figure 5.7

$$\text{AREA} = (15 - 4)\left(\frac{1}{20}\right) = 0.55$$

$$(15 - 4) = 11 = \text{the base of a rectangle}$$

The area corresponds to the probability $P(4 < x < 15) = 0.55$.

Suppose we want to find $P(x = 15)$. On an x-y graph, $x = 15$ is a vertical line. A vertical line has no width (or zero width). Therefore, $P(x = 15)$ = (base)(height) = $(0)\left(\frac{1}{20}\right) = 0$

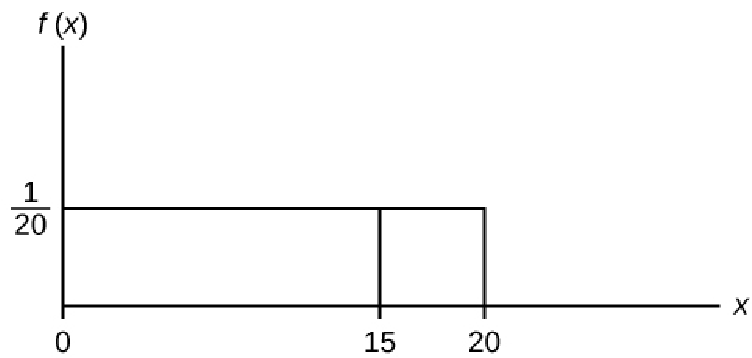



**Figure 5.8**

$P(X \leq x)$, which can also be written as $P(X < x)$ for continuous distributions, is called the cumulative distribution function or CDF. Notice the "less than or equal to" symbol. We can also use the CDF to calculate $P(X > x)$. The CDF gives "area to the left" and $P(X > x)$ gives "area to the right." We calculate $P(X > x)$ for continuous distributions as follows: $P(X > x) = 1 - P(X < x)$.

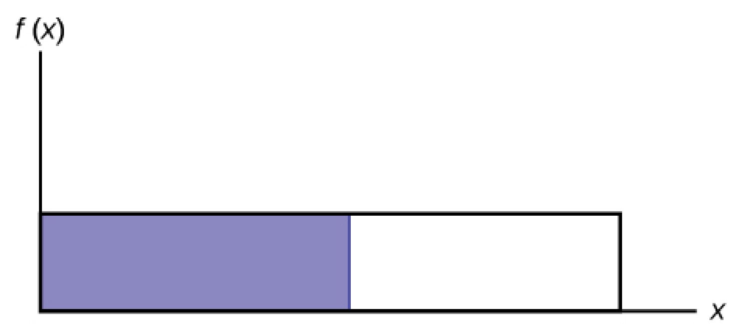



**Figure 5.9**

Label the graph with $f(x)$ and $x$. Scale the $x$ and $y$ axes with the maximum $x$ and $y$ values. $f(x) = \frac{1}{20}$, $0 \le x \le 20$.

To calculate the probability that $x$ is between two values, look at the following graph. Shade the region between $x = 2.3$ and $x = 12.7$. Then calculate the shaded area of a rectangle.

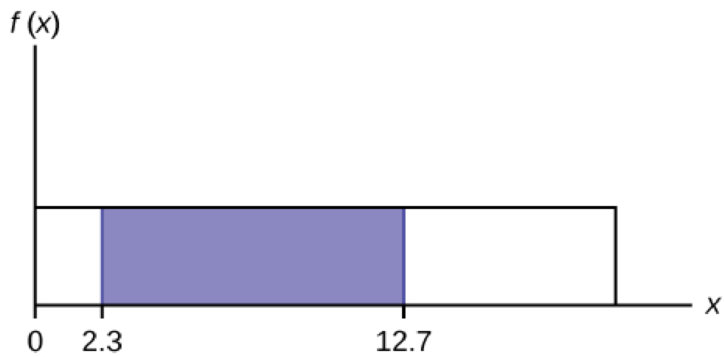



**Figure 5.10**

$$P(2.3 < x < 12.7) = \text{(base)(height)} = (12.7 - 2.3)\left(\frac{1}{20}\right) = 0.52$$

### TRY IT 5.1

Consider the function $f(x) = \frac{1}{8}$ for $0 \le x \le 8$. Draw the graph of $f(x)$ and find $P(2.5 < x < 7.5)$.

The uniform distribution is a continuous probability distribution and is concerned with events that are equally likely to occur. When working out problems that have a uniform distribution, be careful to note if the data is inclusive or exclusive of endpoints.

### EXAMPLE 5.2

The data in Table 5.1 are 55 smiling times, in seconds, of an eight-week-old baby.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 10.4 | 19.6 | 18.8 | 13.9 | 17.8 | 16.8 | 21.6 | 17.9 | 12.5 |
| 12.8 | 14.8 | 22.8 | 20.0 | 15.9 | 16.3 | 13.4 | 17.1 | 14.5 |
| 1.3 | 0.7 | 8.9 | 11.9 | 10.9 | 7.3 | 5.9 | 3.7 | 17.9 |
| 5.8 | 6.9 | 2.6 | 5.8 | 21.7 | 11.8 | 3.4 | 2.1 | 4.5 |
| 8.9 | 9.4 | 9.4 | 7.6 | 10.0 | 3.3 | 6.7 | 7.8 | 11.6 |

**Table 5.1**

The sample mean = 11.49 and the sample standard deviation = 6.23.

We will assume that the smiling times, in seconds, follow a uniform distribution between zero and 23 seconds, inclusive. This means that any smiling time from zero to and including 23 seconds is **equally likely**. The histogram that could be constructed from the sample is an empirical distribution that closely matches the theoretical uniform distribution.

Let $X$ = length, in seconds, of an eight-week-old baby's smile.

The notation for the uniform distribution is

$X \sim U(a, b)$ where $a$ = the lowest value of $x$ and $b$ = the highest value of $x$.

The probability density function is $f(x) = \frac{1}{b-a}$ for $a \le x \le b$.

For this example, $X \sim U(0, 23)$ and $f(x) = \frac{1}{23-0}$ for $0 \le X \le 23$.

Formulas for the theoretical mean and standard deviation are

$$\mu = \frac{a+b}{2} \text{ and } \sigma = \sqrt{\frac{(b-a)^2}{12}}$$

For this problem, the theoretical mean and standard deviation are

$\mu = \frac{0+23}{2}$ = 11.50 seconds and $\sigma = \sqrt{\frac{(23-0)^2}{12}}$ = 6.64 seconds.

Notice that the theoretical mean and standard deviation are close to the sample mean and standard deviation in this example.

## TRY IT 5.2

The data that follow are the number of passengers on 35 different charter fishing boats. The sample mean = 7.9 and the sample standard deviation = 4.33. The data follow a uniform distribution where all values between and including zero and 14 are equally likely. State the values of *a* and *b*. Write the distribution in proper notation, and calculate the theoretical mean and standard deviation.

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | 12 | 4 | 10 | 4 | 14 | 11 |
| 7 | 11 | 4 | 13 | 2 | 4 | 6 |
| 3 | 10 | 0 | 12 | 6 | 9 | 10 |
| 5 | 13 | 4 | 10 | 14 | 12 | 11 |
| 6 | 10 | 11 | 0 | 11 | 13 | 2 |

**Table 5.2**

## EXAMPLE 5.3

a. Refer to Example 5.2. What is the probability that a randomly chosen eight-week-old baby smiles between two and 18 seconds?

**[Show Solution]**

b. Find the 90$^{th}$ percentile for an eight-week-old baby's smiling time.

[Show Solution]

c. Find the probability that a random eight-week-old baby smiles more than 12 seconds **KNOWING** that the baby smiles **MORE THAN EIGHT SECONDS**.

[Show Solution]

### TRY IT 5.3

A distribution is given as $X \sim U(0, 20)$. What is $P(2 < x < 18)$? Find the 90$^{th}$ percentile.

### EXAMPLE 5.4

The amount of time, in minutes, that a person must wait for a bus is uniformly distributed between zero and 15 minutes, inclusive.

a. What is the probability that a person waits fewer than 12.5 minutes?

[Show Solution]

b. On the average, how long must a person wait? Find the mean, $\mu$, and the standard deviation, $\sigma$.

[Show Solution]

c. Ninety percent of the time, the time a person must wait falls below what value?

This asks for the 90$^{th}$ percentile.

**[Show Solution]**

### TRY IT 5.4

The total duration of baseball games in the major league in the 2011 season is uniformly distributed between 447 hours and 521 hours inclusive.

a. Find *a* and *b* and describe what they represent.
b. Write the distribution.
c. Find the mean and the standard deviation.
d. What is the probability that the duration of games for a team for the 2011 season is between 480 and 500 hours?
e. What is the 65$^{th}$ percentile for the duration of games for a team for the 2011 season?

### EXAMPLE 5.5

Suppose the time it takes a nine-year old to eat a donut is between 0.5 and 4 minutes, inclusive. Let *X* = the time, in minutes, it takes a nine-year old child to eat a donut. Then $X \sim U\ (0.5, 4)$.

a. The probability that a randomly selected nine-year old child eats a donut in at least two minutes is _______.

**[Show Solution]**

b. Find the probability that a different nine-year old child eats a donut in more than two minutes given that the child has already been eating the donut for more than 1.5 minutes.

The second question has a **conditional probability**. You are asked to find the probability that a nine-year old child eats a donut in more than two minutes given that the child has already been eating the donut for more than 1.5 minutes. Solve the problem two different ways (see Example 5.3). You must reduce the sample space. **First way**: Since you know the child has already been eating the donut for more than 1.5 minutes, you are no longer starting at *a* = 0.5 minutes. Your starting point is 1.5 minutes.

**Write a new *f*(*x*):**

$f(x) = \frac{1}{4-1.5} = \frac{2}{5}$ for $1.5 \le x \le 4$.

Find $P(x > 2 | x > 1.5)$. Draw a graph.

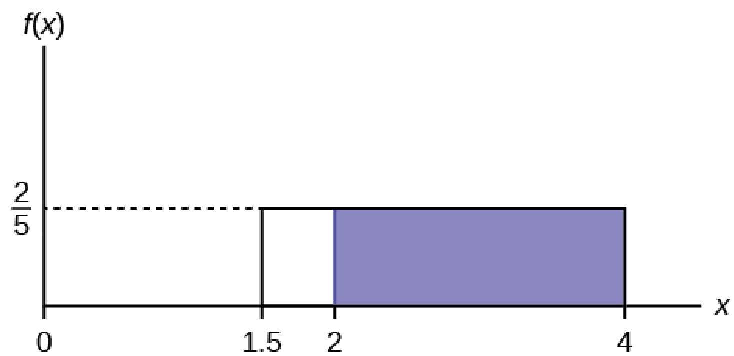



**Figure 5.17**

$P(x > 2 | x > 1.5)$ = (base)(new height) = $(4 - 2)\left(\frac{2}{5}\right) = \frac{4}{5}$

**[Show Solution]**

The probability that a nine-year old child eats a donut in more than two minutes given that the child has already been eating the donut for more than 1.5 minutes is $\frac{4}{5}$.

**Second way:** Draw the original graph for $X \sim U\,(0.5, 4)$. Use the conditional formula

$$P(x > 2 | x > 1.5) = \frac{P(x>2 \text{ AND } x>1.5)}{P(x>1.5)} = \frac{P(x>2)}{P(x>1.5)} = \frac{\frac{2}{3.5}}{\frac{2.5}{3.5}} = 0.8 = \frac{4}{5}$$

## TRY IT 5.5

Suppose the time it takes a student to finish a quiz is uniformly distributed between six and 15 minutes, inclusive. Let $X$ = the time, in minutes, it takes a student to finish a quiz. Then $X \sim U\,(6, 15)$.

Find the probability that a randomly selected student needs at least eight minutes to complete the quiz. Then find the probability that a different student needs at least eight minutes to finish the quiz given that she has already taken more than seven minutes.

### EXAMPLE 5.6

Ace Heating and Air Conditioning Service finds that the amount of time a repairman needs to fix a furnace is uniformly distributed between 1.5 and four hours. Let $x$ = the time needed to fix a furnace. Then $x \sim U(1.5, 4)$.

a. Find the probability that a randomly selected furnace repair requires more than two hours.
b. Find the probability that a randomly selected furnace repair requires less than three hours.
c. Find the 30$^{th}$ percentile of furnace repair times.
d. The longest 25% of furnace repair times take at least how long? (In other words: find the minimum time for the longest 25% of repair times.) What percentile does this represent?
e. Find the mean and standard deviation

[Show Solution]

[Show Solution]

[Show Solution]

[Show Solution]

[Show Solution]

### TRY IT 5.6

The amount of time a service technician needs to change the oil in a car is uniformly distributed between 11 and 21 minutes. Let $X$ = the time needed to change the oil on a car.

a. Write the random variable $X$ in words. $X$ = _________________.
b. Write the distribution.
c. Graph the distribution.
d. Find $P(x > 19)$.
e. Find the $50^{th}$ percentile.

The **exponential distribution** is often concerned with the amount of time until some specific event occurs. For example, the amount of time (beginning now) until an earthquake occurs has an exponential distribution. Other examples include the length, in minutes, of long distance business telephone calls, and the amount of time, in months, a car battery lasts. It can be shown, too, that the value of the change that you have in your pocket or purse approximately follows an exponential distribution.

Values for an exponential random variable occur in the following way. There are fewer large values and more small values. For example, the amount of money customers spend in one trip to the supermarket follows an exponential distribution. There are more people who spend small amounts of money and fewer people who spend large amounts of money.

Exponential distributions are commonly used in calculations of product reliability, or the length of time a product lasts.

### EXAMPLE 5.7

Let *X* = amount of time (in minutes) a postal clerk spends with his or her customer. The time is known to have an exponential distribution with the average amount of time equal to four minutes.

*X* is a **continuous random variable** since time is measured. It is given that $\mu$ = 4 minutes. To do any calculations, you must know *m*, the decay parameter.

$m = \frac{1}{\mu}$. Therefore, $m = \frac{1}{4} = 0.25.$

The standard deviation, $\sigma$, is the same as the mean. $\mu = \sigma$

The distribution notation is *X* ~ *Exp*(*m*). Therefore, *X* ~ *Exp*(0.25).

The probability density function is $f(x) = me^{-mx}$. The number *e* = 2.71828182846... It is a number that is used often in mathematics. Scientific calculators have the key "$e^x$." If you enter one for *x*, the calculator will display the value *e*.

The curve is:

$f(x) = 0.25e^{-0.25x}$ where *x* is at least zero and *m* = 0.25.

For example, $f(5) = 0.25e^{(-0.25)(5)} = 0.072$. The value 0.072 is the height of the curve when *x* = 5. In Example 5.8 below, you will learn how to find probabilities using the decay parameter.

The graph is as follows:

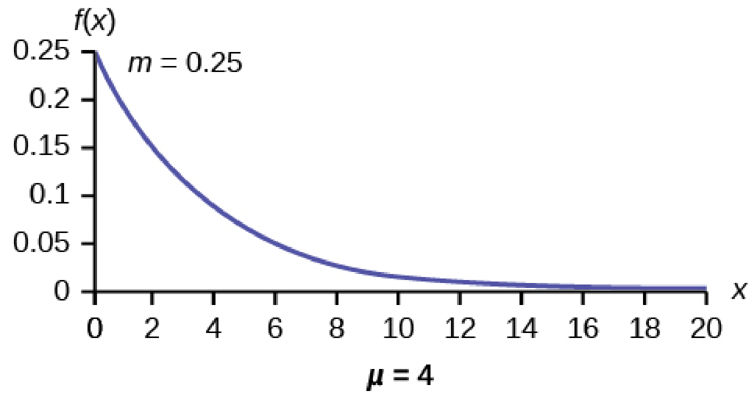



**Figure 5.22**

Notice the graph is a declining curve. When $x = 0$,

$f(x) = 0.25e^{(-0.25)(0)} = (0.25)(1) = 0.25 = m$. The maximum value on the $y$-axis is $m$.

## TRY IT 5.7

The amount of time spouses shop for anniversary cards can be modeled by an exponential distribution with the average amount of time equal to eight minutes. Write the distribution, state the probability density function, and graph the distribution.

## EXAMPLE 5.8

a. Using the information in Example 5.7, find the probability that a clerk spends four to five minutes with a randomly selected customer.

**[Show Solution]**

b. Half of all customers are finished within how long? (Find the 50$^{th}$ percentile)

[Show Solution]

c. Which is larger, the mean or the median?

[Show Solution]

### TRY IT 5.8

The number of days ahead travelers purchase their airline tickets can be modeled by an exponential distribution with the average amount of time equal to 15 days. Find the probability that a traveler will purchase a ticket fewer than ten days in advance. How many days do half of all travelers wait?

### COLLABORATIVE EXERCISE

Have each class member count the change he or she has in his or her pocket or purse. Your instructor will record the amounts in dollars and cents. Construct a histogram of the data taken by the class. Use five intervals. Draw a smooth curve through the bars. The graph should look approximately exponential. Then calculate the mean.

Let $X$ = the amount of money a student in your class has in his or her pocket or purse.

The distribution for $X$ is approximately exponential with mean, $\mu$ = _______ and $m$ = _______. The standard deviation, $\sigma$ = ________.

Draw the appropriate exponential graph. You should label the x– and y–axes, the decay rate, and the mean. Shade the area that represents the probability that one student has less than $.40 in his or her pocket or purse. (Shade $P(x < 0.40)$).

### EXAMPLE 5.9

On the average, a certain computer part lasts ten years. The length of time the computer part lasts is exponentially distributed.

a. What is the probability that a computer part lasts more than 7 years?

[Show Solution]

### USING THE TI-83, 83+, 84, 84+ CALCULATOR

On the home screen, enter e^(-.1*7).

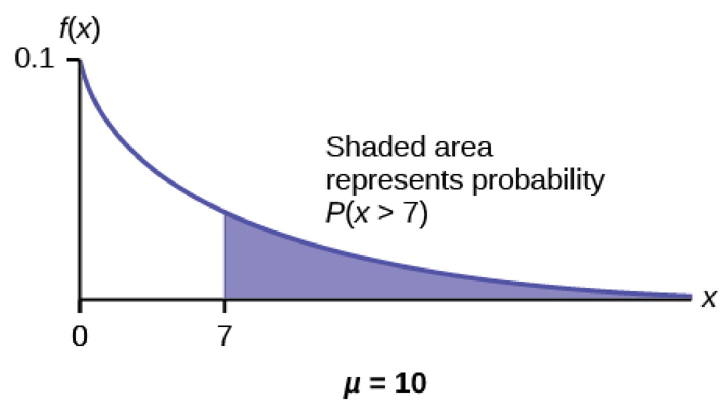



**Figure 5.25**

b. On the average, how long would five computer parts last if they are used one after another?

[Show Solution]

c. Eighty percent of computer parts last at most how long?

[Show Solution]

### USING THE TI-83, 83+, 84, 84+ CALCULATOR

On the home screen, enter $\frac{\ln(1-0.80)}{-0.1}$

d. What is the probability that a computer part lasts between nine and 11 years?

[Show Solution]

### USING THE TI-83, 83+, 84, 84+ CALCULATOR

On the home screen, enter *e*^(–0.1*9) – *e*^(–0.1*11).

### TRY IT 5.9

On average, a pair of running shoes can last 18 months if used every day. The length of time running shoes last is exponentially distributed. What is the probability that a pair of running shoes last more than 15 months? On average, how long would six pairs of running shoes last if they are used one after the other? Eighty percent of running shoes last at most how long if used every day?

### EXAMPLE 5.10

Suppose that the length of a phone call, in minutes, is an exponential random variable with decay parameter $\frac{1}{12}$. If another person arrives at a public telephone just before you, find the probability that you will have to wait more than five minutes. Let $X$ = the length of a phone call, in minutes.

What is *m*, $\mu$, and $\sigma$? The probability that you must wait more than five minutes is _______ .

[Show Solution]

### TRY IT 5.10

Suppose that the distance, in miles, that people are willing to commute to work is an exponential random variable with a decay parameter $\frac{1}{20}$. Let *X* = the distance people are willing to commute in miles. What is *m*, $\mu$, and $\sigma$? What is the probability that a person is willing to commute more than 25 miles?

### EXAMPLE 5.11

The time spent waiting between events is often modeled using the exponential distribution. For example, suppose that an average of 30 customers per hour arrive at a store and the time between arrivals is exponentially distributed.

a. On average, how many minutes elapse between two successive arrivals?
b. When the store first opens, how long on average does it take for three customers to arrive?
c. After a customer arrives, find the probability that it takes less than one minute for the next customer to arrive.
d. After a customer arrives, find the probability that it takes more than five minutes for the next customer to arrive.
e. Seventy percent of the customers arrive within how many minutes of the previous customer?
f. Is an exponential distribution reasonable for this situation?

[Show Solution]

### TRY IT 5.11

Suppose that on a certain stretch of highway, cars pass at an average rate of five cars per minute. Assume that the duration of time between successive cars follows the exponential distribution.

a. On average, how many seconds elapse between two successive cars?
b. After a car passes by, how long on average will it take for another seven cars to pass by?
c. Find the probability that after a car passes by, the next car will pass within the next 20 seconds.
d. Find the probability that after a car passes by, the next car will not pass for at least another 15 seconds.

## Memorylessness of the Exponential Distribution

In Example 5.7 recall that the amount of time between customers is exponentially distributed with a mean of two minutes ($X \sim Exp\ (0.5)$). Suppose that five minutes have elapsed since the last customer arrived. Since an unusually long amount of time has now elapsed, it would seem to be more likely for a customer to arrive within the next minute. With the exponential distribution, this is not the case–the additional time spent waiting for the next customer does not depend on how much time has already elapsed since the last customer. This is referred to as the **memoryless property**. Specifically, the **memoryless property** says that

$P\ (X > r + t \mid X > r) = P\ (X > t)$ for all $r \geq 0$ and $t \geq 0$

For example, if five minutes have elapsed since the last customer arrived, then the probability that more than one minute will elapse before the next customer arrives is computed by using $r = 5$ and $t = 1$ in the foregoing equation.

$P(X > 5 + 1 \mid X > 5) = P(X > 1) = e^{(-0.5)(1)} \approx 0.6065.$

*This is the same* probability as that of waiting more than one minute for a customer to arrive after the previous arrival.

The exponential distribution is often used to model the longevity of an electrical or mechanical device. In Example 5.9, the lifetime of a certain computer part has the exponential distribution with a mean of ten years ($X \sim Exp(0.1)$). The **memoryless property** says that knowledge of what has occurred in the past has no effect on future probabilities. In this case it means that an old part is not any more likely to break down at any particular time than a brand new part. In other words, the part stays as good as new until it suddenly breaks. For example, if the part has already lasted ten years, then the probability that it lasts another seven years is $P(X > 17|X > 10) = P(X > 7) = 0.4966$.

### EXAMPLE 5.12

Refer to Example 5.7 where the time a postal clerk spends with his or her customer has an exponential distribution with a mean of four minutes. Suppose a customer has spent four minutes with a postal clerk. What is the probability that he or she will spend at least an additional three minutes with the postal clerk?

The decay parameter of $X$ is $m = \frac{1}{4} = 0.25$, so $X \sim Exp(0.25)$.

The cumulative distribution function is $P(X < x) = 1 - e^{-0.25x}$.

We want to find $P(X > 7|X > 4)$. The **memoryless property** says that $P(X > 7|X > 4) = P(X > 3)$, so we just need to find the probability that a customer spends more than three minutes with a postal clerk.

This is $P(X > 3) = 1 - P(X < 3) = 1 - (1 - e^{-0.25 \cdot 3}) = e^{-0.75} \approx 0.4724$.

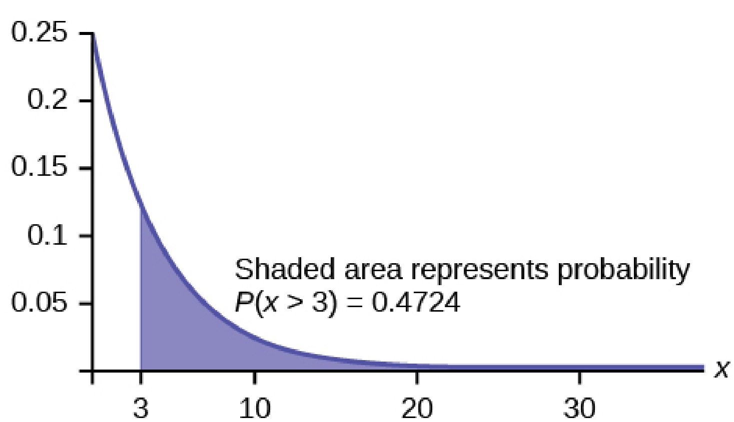



**Figure 5.31**

#### USING THE TI-83, 83+, 84, 84+ CALCULATOR

1–(1–e^(–0.25*3)) = e^(–0.25*3).

### TRY IT 5.12

Suppose that the longevity of a light bulb is exponential with a mean lifetime of eight years. If a bulb has already lasted 12 years, find the probability that it will last a total of over 19 years.

## Relationship between the Poisson and the Exponential Distribution

There is an interesting relationship between the exponential distribution and the Poisson distribution. Suppose that the time that elapses between two successive events follows the exponential distribution with a mean of $\mu$ units of time. Also assume that these times are independent, meaning that the time between events is not affected by the times between previous events. If these assumptions hold, then the number of events per unit time follows a Poisson distribution with mean $\lambda = 1/\mu$. Recall from the chapter on Discrete Random Variables that if $X$ has the Poisson distribution with mean $\lambda$, then $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$. Conversely, if the number of events per unit time follows a Poisson distribution, then the amount of time between events follows the exponential distribution. ($k! = k*(k–1*)(k–2)*(k–3)*…3*2*1$)

### USING THE TI-83, 83+, 84, 84+ CALCULATOR

Suppose $X$ has the Poisson distribution with mean $\lambda$. Compute $P(X = k)$ by entering $2^{nd}$, VARS(DISTR), C: poissonpdf($\lambda$, $k$). To compute $P(X \le k)$, enter $2^{nd}$, VARS (DISTR), D:poissoncdf($\lambda$, $k$).

### EXAMPLE 5.13

At a police station in a large city, calls come in at an average rate of four calls per minute. Assume that the time that elapses from one call to the next has the exponential distribution. Take note that we are concerned only with the rate at which calls come in, and we are ignoring the time spent on the phone. We must also assume that the times spent between calls are independent. This means that a particularly long delay between two calls does not mean that there will be a shorter waiting period for the next call. We may then deduce that the total number of calls received during a time period has the Poisson distribution.

a. Find the average time between two successive calls.
b. Find the probability that after a call is received, the next call occurs in less than ten seconds.
c. Find the probability that exactly five calls occur within a minute.
d. Find the probability that less than five calls occur within a minute.
e. Find the probability that more than 40 calls occur in an eight-minute period.

[Show Solution]

### TRY IT 5.13

In a small city, the number of automobile accidents occur with a Poisson distribution at an average of three per week.

a. Calculate the probability that there are at most 2 accidents occur in any given week.
b. What is the probability that there is at least two weeks between any 2 accidents?

**Figure 6.1** If you ask enough people about their shoe size, you will find that your graphed data is shaped like a bell curve and can be described as normally distributed. (credit: Ömer Ünlü)

## CHAPTER OBJECTIVES

By the end of this chapter, the student should be able to:

- Recognize the normal probability distribution and apply it appropriately.
- Recognize the standard normal probability distribution and apply it appropriately.
- Compare normal probabilities by converting to the standard normal distribution.

The normal, a continuous distribution, is the most important of all the distributions. It is widely used and even more widely abused. Its graph is bell-shaped. You see the bell curve in almost all disciplines. Some of these include psychology, business, economics, the sciences, nursing, and, of course, mathematics. Some of your instructors may use the normal distribution to help determine your grade. Most IQ scores are normally distributed. Often real-estate prices fit a normal distribution. The normal distribution is extremely important, but it cannot be applied to everything in the real world.

In this chapter, you will study the normal distribution, the standard normal distribution, and applications associated with them.

The normal distribution has two parameters (two numerical descriptive measures): the mean ($\mu$) and the standard deviation ($\sigma$). If $X$ is a quantity to be measured that has a normal distribution with mean ($\mu$) and standard deviation ($\sigma$), we designate this by writing

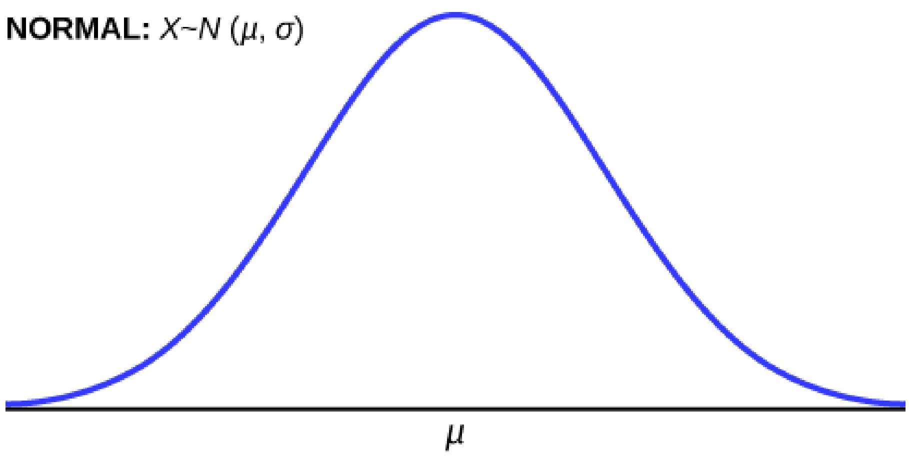



**Figure 6.2**

The probability density function is a rather complicated function. **Do not memorize it**. It is not necessary.

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

The cumulative distribution function is $P(X < x)$. It is calculated either by a calculator or a computer, or it is looked up in a table. Technology has made the tables virtually obsolete. For that reason, as well as the fact that there are various table formats, we are not including table instructions.

The curve is symmetric about a vertical line drawn through the mean, $\mu$. In theory, the mean is the same as the median, because the graph is symmetric about $\mu$. As the notation indicates, the normal distribution depends only on the mean and the standard deviation. Since the area under the curve must equal one, a change in the standard deviation, $\sigma$, causes a change in the shape of the curve; the curve becomes fatter or skinnier depending on $\sigma$. A change in $\mu$ causes the graph to shift to the left or right. This means there are an infinite number of normal probability distributions. One of special interest is called the **standard normal distribution**.

**COLLABORATIVE EXERCISE**

Your instructor will record the heights of both men and women in your class, separately. Draw histograms of your data. Then draw a smooth curve through each histogram. Is each curve somewhat bell-shaped? Do you think that if you had recorded 200 data values for men and 200 for women that the curves would look bell-shaped? Calculate the mean for each data set. Write the means on the *x*-axis of the appropriate graph below the peak. Shade the approximate area that represents the probability that one randomly chosen male is taller than 72 inches. Shade the approximate area that represents the probability that one randomly chosen female is shorter than 60 inches. If the total area under each curve is one, does either probability appear to be more than 0.5?

The **standard normal distribution** is a normal distribution of **standardized values called** ***z*-scores**. **A *z*-score is measured in units of the standard deviation.** For example, if the mean of a normal distribution is five and the standard deviation is two, the value 11 is three standard deviations above (or to the right of) the mean. The calculation is as follows:

$x = \mu + (z)(\sigma) = 5 + (3)(2) = 11$

The *z*-score is three.

The mean for the standard normal distribution is zero, and the standard deviation is one. The transformation $z = \frac{x-\mu}{\sigma}$ produces the distribution $Z \sim N(0, 1)$. The value $x$ in the given equation comes from a normal distribution with mean $\mu$ and standard deviation $\sigma$.

## *Z*-Scores

If $X$ is a normally distributed random variable and $X \sim N(\mu, \sigma)$, then the *z*-score is:

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

**The *z*-score tells you how many standard deviations the value *x* is above (to the right of) or below (to the left of) the mean, $\mu$.** Values of $x$ that are larger than the mean have positive *z*-scores, and values of $x$ that are smaller than the mean have negative *z*-scores. If $x$ equals the mean, then $x$ has a *z*-score of zero.

### EXAMPLE 6.1

Suppose $X \sim N(5, 6)$. This says that $X$ is a normally distributed random variable with mean $\mu = 5$ and standard deviation $\sigma = 6$. Suppose $x = 17$. Then:

$$z = \frac{x-\mu}{\sigma} = \frac{17-5}{6} = 2$$

This means that $x = 17$ is **two standard deviations** ($2\sigma$) above or to the right of the mean $\mu = 5$.

Notice that: 5 + (2)(6) = 17 (The pattern is $\mu + z\sigma = x$)

Now suppose $x = 1$. Then: $z = \frac{x-\mu}{\sigma} = \frac{1-5}{6} = -0.67$ (rounded to two decimal places)

**This means that $x = 1$ is 0.67 standard deviations ($-0.67\sigma$) below or to the left of the mean $\mu = 5$. Notice that:** 5 + (–0.67)(6) is approximately equal to one (This has the pattern $\mu + (-0.67)\sigma = 1$)

Summarizing, when $z$ is positive, $x$ is above or to the right of $\mu$ and when $z$ is negative, $x$ is to the left of or below $\mu$. Or, when $z$ is positive, $x$ is greater than $\mu$, and when $z$ is negative $x$ is less than $\mu$.

### TRY IT 6.1

What is the $z$-score of $x$, when $x = 1$ and $X \sim N(12,3)$?

### EXAMPLE 6.2

Some doctors believe that a person can lose five pounds, on the average, in a month by reducing his or her fat intake and by exercising consistently. Suppose weight loss has a normal distribution. Let $X$ = the amount of weight lost (in pounds) by a person in a month. Use a standard deviation of two pounds. $X \sim N(5, 2)$. Fill in the blanks.

a. Suppose a person **lost** ten pounds in a month. The $z$-score when $x = 10$ pounds is $z = 2.5$ (verify). This $z$-score tells you that $x = 10$ is ________ standard deviations to the ________ (right or left) of the mean _____ (What is the mean?).

[Show Solution]

b. Suppose a person **gained** three pounds (a negative weight loss). Then $z$ = __________. This $z$-score tells you that $x = -3$ is ________ standard deviations to the __________ (right or left) of the mean.

[Show Solution]

c. Suppose the random variables $X$ and $Y$ have the following normal distributions: $X \sim N(5, 6)$ and $Y \sim N(2, 1)$. If $x = 17$, then $z = 2$. (This was previously shown.) If $y = 4$, what is $z$?

[Show Solution]

The $z$-score for $y = 4$ is $z = 2$. This means that four is $z = 2$ standard deviations to the right of the mean. Therefore, $x = 17$ and $y = 4$ are both two (of **their own**) standard deviations to the right of **their** respective means.

**The *z*-score allows us to compare data that are scaled differently.** To understand the concept, suppose $X \sim N(5, 6)$ represents weight gains for one group of people who are trying to gain weight in a six week period and $Y \sim N(2, 1)$ measures the same weight gain for a second group of people. A negative weight gain would be a weight loss. Since $x = 17$ and $y = 4$ are each two standard deviations to the right of their means, they represent the same, standardized weight gain **relative to their means**.

**TRY IT 6.2**

Fill in the blanks.

Jerome averages 16 points a game with a standard deviation of four points. $X \sim N(16,4)$. Suppose Jerome scores ten points in a game. The $z$–score when $x = 10$ is –1.5. This score tells you that $x = 10$ is _____ standard deviations to the ______(right or left) of the mean______(What is the mean?).

**The Empirical Rule**

If $X$ is a random variable and has a normal distribution with mean $\mu$ and standard deviation $\sigma$, then the **Empirical Rule** states the following:

- About 68% of the $x$ values lie between $-1\sigma$ and $+1\sigma$ of the mean $\mu$ (within one standard deviation of the mean).
- About 95% of the $x$ values lie between $-2\sigma$ and $+2\sigma$ of the mean $\mu$ (within two standard deviations of the mean).
- About 99.7% of the $x$ values lie between $-3\sigma$ and $+3\sigma$ of the mean $\mu$ (within three standard deviations of the mean). Notice that almost all the $x$ values lie within three standard deviations of the mean.
- The $z$-scores for $+1\sigma$ and $-1\sigma$ are +1 and –1, respectively.
- The $z$-scores for $+2\sigma$ and $-2\sigma$ are +2 and –2, respectively.
- The $z$-scores for $+3\sigma$ and $-3\sigma$ are +3 and –3 respectively.

The empirical rule is also known as the 68-95-99.7 rule.

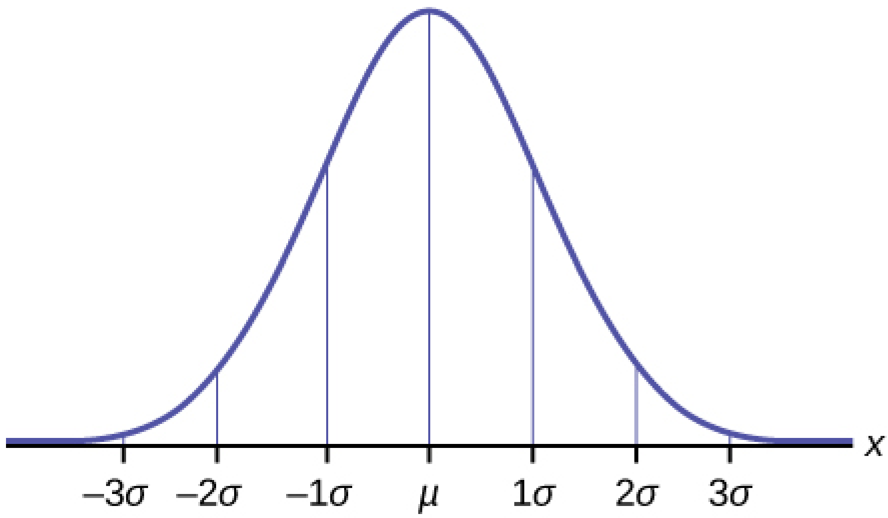



**Figure 6.3**

## EXAMPLE 6.3

The mean height of 15 to 18-year-old males from Chile from 2009 to 2010 was 170 cm with a standard deviation of 6.28 cm. Male heights are known to follow a normal distribution. Let $X$ = the height of a 15 to 18-year-old male from Chile in 2009 to 2010. Then $X \sim N(170, 6.28)$.

a. Suppose a 15 to 18-year-old male from Chile was 168 cm tall from 2009 to 2010. The *z*-score when $x$ = 168 cm is $z$ = _______. This *z*-score tells you that $x$ = 168 is ________ standard deviations to the ________ (right or left) of the mean _____ (What is the mean?).

**[Show Solution]**

b. Suppose that the height of a 15 to 18-year-old male from Chile from 2009 to 2010 has a *z*-score of $z$ = 1.27. What is the male's height? The *z*-score ($z$ = 1.27) tells you that the male's height is ________ standard deviations to the __________ (right or left) of the mean.

**[Show Solution]**

## TRY IT 6.3

Use the information in Example 6.3 to answer the following questions.

a. Suppose a 15 to 18-year-old male from Chile was 176 cm tall from 2009 to 2010. The $z$-score when $x$ = 176 cm is $z$ = _______. This $z$-score tells you that $x$ = 176 cm is ________ standard deviations to the ________ (right or left) of the mean _____ (What is the mean?).
b. Suppose that the height of a 15 to 18-year-old male from Chile from 2009 to 2010 has a $z$-score of $z$ = –2. What is the male's height? The $z$-score ($z$ = –2) tells you that the male's height is ________ standard deviations to the __________ (right or left) of the mean.

## EXAMPLE 6.4

From 1984 to 1985, the mean height of 15 to 18-year-old males from Chile was 172.36 cm, and the standard deviation was 6.34 cm. Let $Y$ = the height of 15 to 18-year-old males from 1984 to 1985. Then $Y \sim N(172.36, 6.34)$.

The mean height of 15 to 18-year-old males from Chile from 2009 to 2010 was 170 cm with a standard deviation of 6.28 cm. Male heights are known to follow a normal distribution. Let $X$ = the height of a 15 to 18-year-old male from Chile in 2009 to 2010. Then $X \sim N(170, 6.28)$.

Find the $z$-scores for $x$ = 160.58 cm and $y$ = 162.85 cm. Interpret each $z$-score. What can you say about $x$ = 160.58 cm and $y$ = 162.85 cm as they compare to their respective means and standard deviations?

**[Show Solution]**

## TRY IT 6.4

In 2012, 1,664,479 students took the SAT exam. The distribution of scores in the verbal section of the SAT had a mean $\mu$ = 496 and a standard deviation $\sigma$ = 114. Let $X$ = a SAT exam verbal section score in 2012. Then $X \sim N(496, 114)$.

Find the $z$-scores for $x_1 = 325$ and $x_2 = 366.21$. Interpret each $z$-score. What can you say about $x_1 = 325$ and $x_2 = 366.21$ as they compare to their respective means and standard deviations?

### EXAMPLE 6.5

Suppose $x$ has a normal distribution with mean 50 and standard deviation 6.

- About 68% of the $x$ values lie within one standard deviation of the mean. Therefore, about 68% of the $x$ values lie between $-1\sigma = (-1)(6) = -6$ and $1\sigma = (1)(6) = 6$ of the mean 50. The values $50 - 6 = 44$ and $50 + 6 = 56$ are within one standard deviation from the mean 50. The $z$-scores are –1 and +1 for 44 and 56, respectively.
- About 95% of the $x$ values lie within two standard deviations of the mean. Therefore, about 95% of the $x$ values lie between $-2\sigma = (-2)(6) = -12$ and $2\sigma = (2)(6) = 12$. The values $50 - 12 = 38$ and $50 + 12 = 62$ are within two standard deviations from the mean 50. The $z$-scores are –2 and +2 for 38 and 62, respectively.
- About 99.7% of the $x$ values lie within three standard deviations of the mean. Therefore, about 99.7% of the $x$ values lie between $-3\sigma = (-3)(6) = -18$ and $3\sigma = (3)(6) = 18$ from the mean 50. The values $50 - 18 = 32$ and $50 + 18 = 68$ are within three standard deviations of the mean 50. The $z$-scores are –3 and +3 for 32 and 68, respectively.

### TRY IT 6.5

Suppose $X$ has a normal distribution with mean 25 and standard deviation five. Between what values of $x$ do 68% of the values lie?

### EXAMPLE 6.6

From 1984 to 1985, the mean height of 15 to 18-year-old males from Chile was 172.36 cm, and the standard deviation was 6.34 cm. Let $Y$ = the height of 15 to 18-year-old males in 1984 to 1985. Then $Y \sim N(172.36, 6.34)$.

a. About 68% of the $y$ values lie between what two values? These values are ________________. The $z$-scores are ________________, respectively.
b. About 95% of the $y$ values lie between what two values? These values are ________________. The $z$-scores are ________________ respectively.
c. About 99.7% of the $y$ values lie between what two values? These values are ________________. The $z$-scores are ________________, respectively.

[Show Solution]

### TRY IT 6.6

The scores on a college entrance exam have an approximate normal distribution with mean, $\mu$ = 52 points and a standard deviation, $\sigma$ = 11 points.

a. About 68% of the $y$ values lie between what two values? These values are ________________. The $z$-scores are ________________, respectively.
b. About 95% of the $y$ values lie between what two values? These values are ________________. The $z$-scores are ________________, respectively.
c. About 99.7% of the $y$ values lie between what two values? These values are ________________. The $z$-scores are ________________, respectively.

The shaded area in the following graph indicates the area to the left of $x$. This area is represented by the probability $P(X < x)$. Normal tables, computers, and calculators provide or calculate the probability $P(X < x)$.

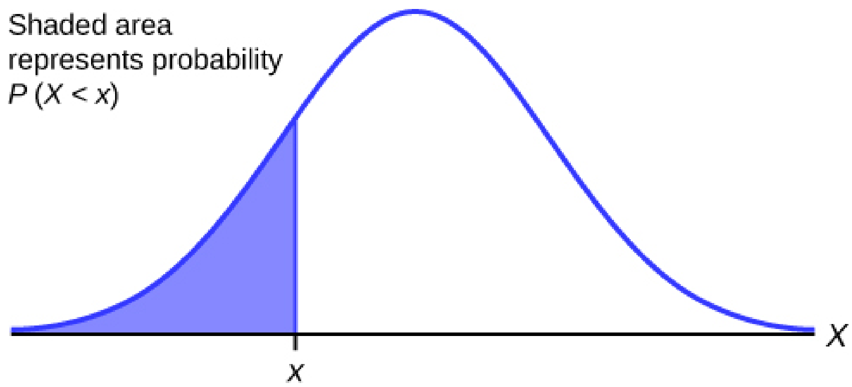



**Figure 6.4**

The area to the right is then $P(X > x) = 1 - P(X < x)$. Remember, $P(X < x)$ = **Area to the left** of the vertical line through $x$ . $P(X > x) = 1 - P(X < x)$ = **Area to the right** of the vertical line through $x$ . $P(X < x)$ is the same as $P(X \leq x)$ and $P(X > x)$ is the same as $P(X \geq x)$ for continuous distributions.

## Calculations of Probabilities

Probabilities are calculated using technology. There are instructions given as necessary for the TI-83+ and TI-84 calculators.

**NOTE**

To calculate the probability, use the probability tables provided in Figure G1 without the use of technology. The tables include instructions for how to use them.

**EXAMPLE 6.7**

If the area to the left is 0.0228, then the area to the right is 1 – 0.0228 = 0.9772.

### TRY IT 6.7

If the area to the left of $x$ is 0.012, then what is the area to the right?

### EXAMPLE 6.8

The final exam scores in a statistics class were normally distributed with a mean of 63 and a standard deviation of five.

a. Find the probability that a randomly selected student scored more than 65 on the exam.

[Show Solution]

#### USING THE TI-83, 83+, 84, 84+ CALCULATOR

Find the percentile for a student scoring 65:

*Press `2nd Distr`
*Press `2:normalcdf`(
*Enter lower bound, upper bound, mean, standard deviation followed by )
*Press `ENTER`.
For this Example, the steps are
`2nd Distr`
`2:normalcdf`(65,1,2nd EE,99,63,5) `ENTER`

The probability that a selected student scored more than 65 is 0.3446.
To find the probability that a selected student scored *more than* 65, subtract the percentile from 1.

b. Find the probability that a randomly selected student scored less than 85.

[Show Solution]

c. Find the 90th percentile (that is, find the score *k* that has 90% of the scores below *k* and 10% of the scores above *k*).

[Show Solution]

#### USING THE TI-83, 83+, 84, 84+ CALCULATOR

`invNorm` in `2nd DISTR`. invNorm(area to the left, mean, standard deviation)
For this problem, invNorm(0.90,63,5) = 69.4

d. Find the 70th percentile (that is, find the score *k* such that 70% of scores are below *k* and 30% of the scores are above *k*).

[Show Solution]

### TRY IT 6.8

The golf scores for a school team were normally distributed with a mean of 68 and a standard deviation of three.

Find the probability that a randomly selected golfer scored less than 65.

### EXAMPLE 6.9

A personal computer is used for office work at home, research, communication, personal finances, education, entertainment, social networking, and a myriad of other things. Suppose that the average number of hours a household personal computer is used for entertainment is two hours per day. Assume the times for entertainment are normally distributed and the standard deviation for the times is half an hour.

a. Find the probability that a household personal computer is used for entertainment between 1.8 and 2.75 hours per day.

**[Show Solution]**

b. Find the maximum number of hours per day that the bottom quartile of households uses a personal computer for entertainment.

**[Show Solution]**

### TRY IT 6.9

The golf scores for a school team were normally distributed with a mean of 68 and a standard deviation of three. Find the probability that a golfer scored between 66 and 70.

### EXAMPLE 6.10

In the United States the ages 13 to 55+ of smartphone users approximately follow a normal distribution with approximate mean and standard deviation of 36.9 years and 13.9 years, respectively.

a. Determine the probability that a random smartphone user in the age range 13 to 55+ is between 23 and 64.7 years old.

**[Show Solution]**

b. Determine the probability that a randomly selected smartphone user in the age range 13 to 55+ is at most 50.8 years old.

**[Show Solution]**

c. Find the 80th percentile of this distribution, and interpret it in a complete sentence.

**[Show Solution]**

### TRY IT 6.10

Use the information in Example 6.10 to answer the following questions.

a. Find the 30th percentile, and interpret it in a complete sentence.
b. What is the probability that the age of a randomly selected smartphone user in the range 13 to 55+ is less than 27 years old.

### EXAMPLE 6.11

In the United States the ages 13 to 55+ of smartphone users approximately follow a normal distribution with approximate mean and standard deviation of 36.9 years and 13.9 years respectively. Using this information, answer the following questions (round answers to one decimal place).

a. Calculate the interquartile range (*IQR*).

**[Show Solution]**

b. Forty percent of the smartphone users from 13 to 55+ are at least what age?

**[Show Solution]**

### TRY IT 6.11

Two thousand students took an exam. The scores on the exam have an approximate normal distribution with a mean $\mu = 81$ points and standard deviation $\sigma = 15$ points.

a. Calculate the first- and third-quartile scores for this exam.
b. The middle 50% of the exam scores are between what two values?

## EXAMPLE 6.12

A citrus farmer who grows mandarin oranges finds that the diameters of mandarin oranges harvested on his farm follow a normal distribution with a mean diameter of 5.85 cm and a standard deviation of 0.24 cm.

a. Find the probability that a randomly selected mandarin orange from this farm has a diameter larger than 6.0 cm. Sketch the graph.

**[Show Solution]**

b. The middle 20% of mandarin oranges from this farm have diameters between ______ and ______.

**[Show Solution]**

c. Find the 90$^{th}$ percentile for the diameters of mandarin oranges, and interpret it in a complete sentence.

**[Show Solution]**

## TRY IT 6.12

Using the information from Example 6.12, answer the following:

a. The middle 40% of mandarin oranges from this farm are between ______ and ______.
b. Find the 16$^{th}$ percentile and interpret it in a complete sentence.