Estimating with Upper and Lower Sums

The *definite integral* is the key tool in calculus for defining and calculating quantities important to mathematics and science, such as areas, volumes, lengths of curved paths, probabilities, and the weights of various objects, just to mention a few. The idea behind the integral is that we can effectively compute such quantities by breaking them into small pieces and then summing the contributions from each piece. We then consider what happens when more and more, smaller and smaller pieces are taken in the summation process. Finally, if the number of terms contributing to the sum approaches infinity and we take the limit of these sums in the way described in Section 5.3, the result is a definite integral. We prove in Section 5.4 that integrals are connected to antiderivatives, a connection that is one of the most important relationships in calculus.

The basis for formulating definite integrals is the construction of appropriate finite sums. Although we need to define precisely what we mean by the area of a general region in the plane, or the average value of a function over a closed interval, we do have intuitive ideas of what these notions mean. So in this section we begin our approach to integration by *approximating* these quantities with finite sums. We also consider what happens when we take more and more terms in the summation process. In subsequent sections we look at taking the limit of these sums as the number of terms goes to infinity, which then leads to precise definitions of the quantities being approximated here.

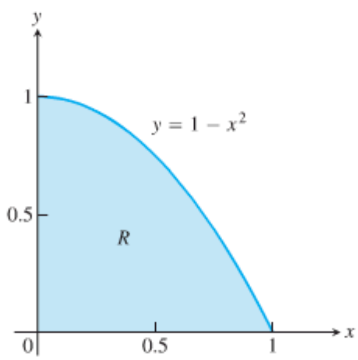


FIGURE 5.1 The area of the region R cannot be found by a simple formula.

Area

Suppose we want to find the area of the shaded region R that lies above the x-axis, below the graph of $y = 1 - x^2$, and between the vertical lines $x = 0$ and $x = 1$ (Figure 5.1). Unfortunately, there is no simple geometric formula for calculating the areas of general shapes having curved boundaries like the region R. How, then, can we find the area of R?

While we do not yet have a method for determining the exact area of R, we can approximate it in a simple way. Figure 5.2a shows two rectangles that together contain the region R. Each rectangle has width $1/2$ and they have heights 1 and $3/4$, moving from left to right. The height of each rectangle is the maximum value of the function f,

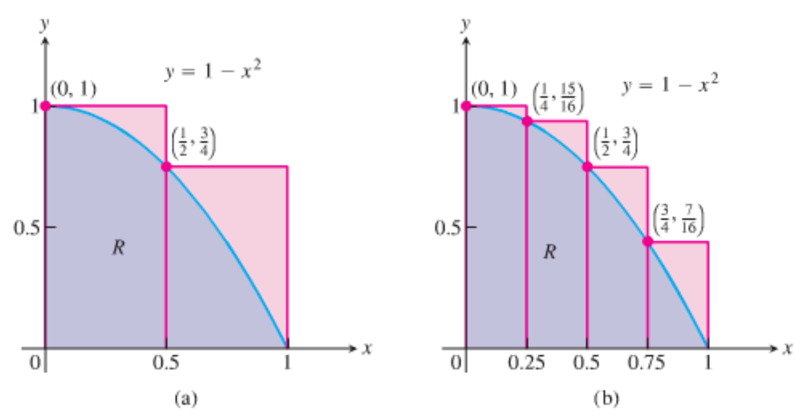


FIGURE 5.2 (a) We get an upper estimate of the area of R by using two rectangles containing R. (b) Four rectangles give a better upper estimate. Both estimates overshoot the true value for the area by the amount shaded in light red.

obtained by evaluating f at the left endpoint of the subinterval of $[0, 1]$ forming the base of the rectangle. The total area of the two rectangles approximates the area A of the region R,

$$A \approx 1 \cdot \frac{1}{2} + \frac{3}{4} \cdot \frac{1}{2} = \frac{7}{8} = 0.875.$$

This estimate is larger than the true area A since the two rectangles contain R. We say that 0.875 is an **upper sum** because it is obtained by taking the height of each rectangle as the maximum (uppermost) value of $f(x)$ for a point x in the base interval of the rectangle. In Figure 5.2b, we improve our estimate by using four thinner rectangles, each of width $1/4$, which taken together contain the region R. These four rectangles give the approximation

$$A \approx 1 \cdot \frac{1}{4} + \frac{15}{16} \cdot \frac{1}{4} + \frac{3}{4} \cdot \frac{1}{4} + \frac{7}{16} \cdot \frac{1}{4} = \frac{25}{32} = 0.78125,$$

which is still greater than A since the four rectangles contain R.

Suppose instead we use four rectangles contained *inside* the region R to estimate the area, as in Figure 5.3a. Each rectangle has width $1/4$ as before, but the rectangles are

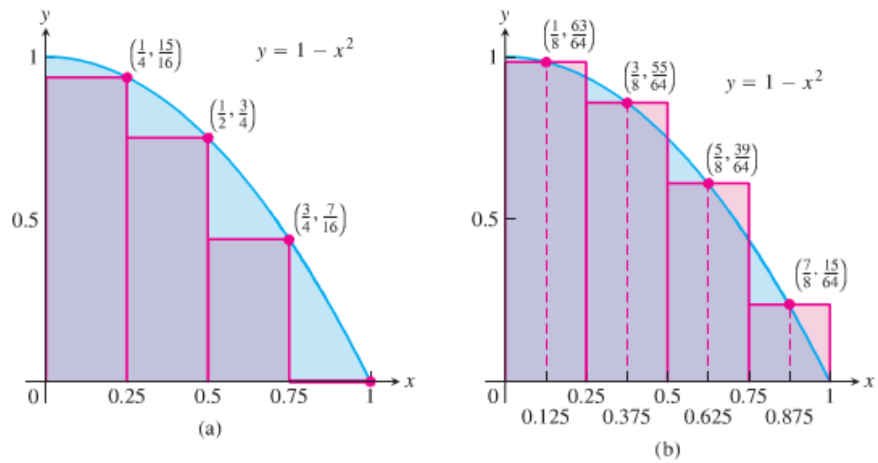


FIGURE 5.3 (a) Rectangles contained in R give an estimate for the area that undershoots the true value by the amount shaded in light blue. (b) The midpoint rule uses rectangles whose height is the value of $y = f(x)$ at the midpoints of their bases. The estimate appears closer to the true value of the area because the light red overshoot areas roughly balance the light blue undershoot areas.

shorter and lie entirely beneath the graph of f. The function $f(x) = 1 - x^2$ is decreasing on $[0, 1]$, so the height of each of these rectangles is given by the value of f at the right endpoint of the subinterval forming its base. The fourth rectangle has zero height and therefore contributes no area. Summing these rectangles with heights equal to the minimum value of $f(x)$ for a point x in each base subinterval gives a **lower sum** approximation to the area,

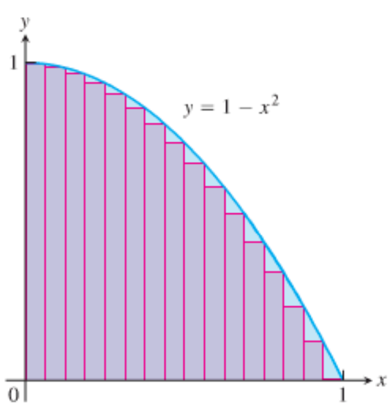
$$A \approx \frac{15}{16} \cdot \frac{1}{4} + \frac{3}{4} \cdot \frac{1}{4} + \frac{7}{16} \cdot \frac{1}{4} + 0 \cdot \frac{1}{4} = \frac{17}{32} = 0.53125.$$

This estimate is smaller than the area A since the rectangles all lie inside of the region R. The true value of A lies somewhere between these lower and upper sums:

$$0.53125 < A < 0.78125.$$

By considering both lower and upper sum approximations we get not only estimates for the area, but also a bound on the size of the possible error in these estimates since the true value of the area lies somewhere between them. Here the error cannot be greater than the difference $0.78125 - 0.53125 = 0.25$.

Yet another estimate can be obtained by using rectangles whose heights are the values of f at the midpoints of their bases (Figure 5.3b). This method of estimation is called the **midpoint rule** for approximating the area. The midpoint rule gives an estimate that is between a lower sum and an upper sum, but it is not quite so clear whether it overestimates or underestimates the true area. With four rectangles of width $1/4$ as before, the midpoint rule estimates the area of R to be



(a)

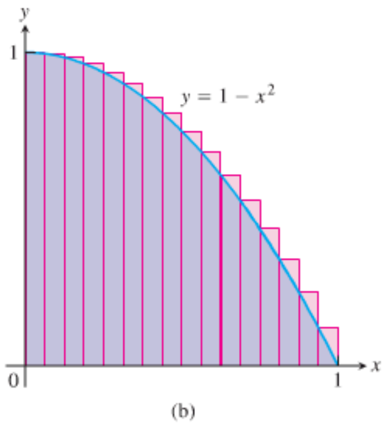


FIGURE 5.4 (a) A lower sum using 16 rectangles of equal width $\Delta x = 1/16$. (b) An upper sum using 16 rectangles.

$$A \approx \frac{63}{64} \cdot \frac{1}{4} + \frac{55}{64} \cdot \frac{1}{4} + \frac{39}{64} \cdot \frac{1}{4} + \frac{15}{64} \cdot \frac{1}{4} = \frac{172}{64} \cdot \frac{1}{4} = 0.671875.$$

In each of our computed sums, the interval $[a, b]$ over which the function f is defined was subdivided into n subintervals of equal width (also called length) $\Delta x = (b - a)/n$, and f was evaluated at a point in each subinterval: c_1 in the first subinterval, c_2 in the second subinterval, and so on. The finite sums then all take the form

$$f(c_1)\,\Delta x + f(c_2)\,\Delta x + f(c_3)\,\Delta x + \cdots + f(c_n)\,\Delta x.$$

By taking more and more rectangles, with each rectangle thinner than before, it appears that these finite sums give better and better approximations to the true area of the region R.

Figure 5.4a shows a lower sum approximation for the area of R using 16 rectangles of equal width. The sum of their areas is 0.634765625, which appears close to the true area, but is still smaller since the rectangles lie inside R.

Figure 5.4b shows an upper sum approximation using 16 rectangles of equal width. The sum of their areas is 0.697265625, which is somewhat larger than the true area because the rectangles taken together contain R. The midpoint rule for 16 rectangles gives a total area approximation of 0.6669921875, but it is not immediately clear whether this estimate is larger or smaller than the true area.