11
More Applications of Integration

11.1 Center of Mass

Suppose a beam is 10 meters long, and that there are three weights on the beam: a 10 kilogram weight at 3 meters from the left end, a 5 kilogram weight 6 meters from the left end, and a 4 kilogram weight 8 meters from the left end. Where should a fulcrum be placed so that the beam balances? Let’s assign a scale to the beam, from 0 at the left end to 10 at the right, so that we can denote locations on the beam simply as $x$ coordinates; the weights are at $x = 3, x = 6,$ and $x = 8$, as in figure 11.1.1.

$$\begin{array}{ccc} 3 & \Delta & 6 \\ 4 & \end{array}$$

Figure 11.1.1 A beam with three masses.

Suppose to begin with that the fulcrum is placed at $x = 5$. What will happen? Each weight applies a force to the beam that tends to rotate it around the fulcrum; this effect is measured by a quantity called torque, proportional to the mass times the distance from the fulcrum. Of course, weights on different sides of the fulcrum rotate the beam in opposite directions. We can distinguish this by using a signed distance in the formula for torque. So with the fulcrum at 5, the torques induced by the three weights will be proportional to $(3 - 5)10 = -20, (6 - 5)5 = 5,$ and $(8 - 5)4 = 12.$ For the beam to balance, the sum of the torques must be zero; since the sum is $-20 + 5 + 12 = -3$, the beam rotates counter-clockwise, and to get the beam to balance we need to move the fulcrum to the left. To calculate exactly where the fulcrum should be, we let $\bar{x}$ denote the location of the fulcrum when the beam is in balance. The total torque on the beam is then $(3 - 5)10 + (6 - 5)5 + (8 - 5)4 = 92 - 19$. Since the beam balances at $\bar{x}$ it must be that $92 - 19 = 0$ or $\bar{x} = 92/19 \approx 4.84$, that is, the fulcrum should be placed at $x = 92/19$ to balance the beam.

Now suppose that we have a beam with varying density—some portions of the beam contain more mass than other portions of the same size. We want to figure out where to put the fulcrum so that the beam balances.

EXAMPLE 11.1.1 Suppose the beam is 10 meters long and that the density is $1 + x$ kilograms per meter at location $x$ on the beam. To approximate the solution, we can think of the beam as a sequence of weights "on" a beam. For example, we can think of the portion of the beam between $x = 0$ and $x = 1$ as a weight sitting at $x = 0$, the portion between $x = 1$ and $x = 2$ as a weight sitting at $x = 1$, and so on, as indicated in figure 11.1.2. We then approximate the mass of the weights by assuming that each portion of the beam has constant density. So the mass of the first weight is approximately $m_0 = (1 + 0)1 = 1$ kilograms, namely, $(1 + 0)$ kilograms per meter times 1 meter. The second weight is $m_1 = (1 + 1)1 = 2$ kilograms, and so on to the tenth weight with $m_9 = (1 + 9)1 = 10$ kilograms. So in this case the total torque is

$$\int_0^1 (0 - x)m_0 + (1 - x)m_1 + \cdots + (9 - x)m_9 = (0 - x)1 + (1 - x)2 + \cdots + (9 - x)10$$

If we set this to zero and solve for $\bar{x}$ we get $\bar{x} = 6$. In general, if we divide the beam into $n$ portions, the mass of weight number $i$ will be $m_i = (1 + i)_i(x_{i+1} - x_i) = (1 + x_i)\Delta x$ and the torque induced by weight number $i$ will be $(x_i - \bar{x})m_i = (x_i - \bar{x})(1 + x_i)\Delta x$. The total torque is then

$$(x_0 - \bar{x})(1 + x_0)\Delta x + (x_1 - \bar{x})(1 + x_1)\Delta x + \cdots + (x_{n-1} - \bar{x})(1 + x_{n-1})\Delta x$$

and the center of mass is at $\bar{x} = \frac{\sum_{i=0}^{n-1} x_i (1 + x_i)\Delta x}{\sum_{i=0}^{n-1} (1 + x_i)\Delta x}$. The numerator of this fraction has a very familiar interpretation. Consider one term of the sum in the denominator: $(1 + x_i)\Delta x$. This is the density near $x_i$ times a short length, $\Delta x$, which in other words is approximately the mass of the beam between $x_i$ and $x_{i+1}$. When we add these up we get approximately the mass of the beam.

Now each of the sums in the fraction has the right form to turn into an integral, which in turn gives us the exact value of $\bar{x}$:

$$\bar{x} = \frac{\int_0^1 x(1 + x)dx}{\int_0^1 (1 + x)dx}.$$  

The numerator of this fraction is called the moment of the system around zero:

$$\int_0^1 x(1 + x)dx = \int_0^1 x^2 + x dx = \frac{1150}{3} +$$

and the denominator is the mass of the beam:

$$\int_0^1 (1 + x) dx = 60,$$

and the balance point, officially called the center of mass, is

$$\bar{x} = \frac{1150}{4} = 135 \approx 6.39.$$
Compute the moment around the $x$-axis:

$$M_y = \int_1^4 y(1 - x^2) \, dx = \frac{1}{4}$$

and the total mass

$$M = \int_1^4 (1 - x^2) \, dx = \frac{2}{3}$$

and finally

$$\bar{y} = \frac{13}{42} \approx \frac{1}{3}.$$
11.2 Kinetic energy; improper integrals

Now suppose that the initial velocity of the object, \( v_0 \), is just enough to get it to infinity, that is, just enough so that the object never slows to a stop, but so that its speed decreases to zero, i.e., so that \( v(t) = 0 \). Then

\[
\frac{m v^2}{2} - \frac{m v_0^2}{2} = -\frac{m g h}{2}
\]

so

\[
v_0 = \sqrt{\frac{2gh}{m}} = 11181 \text{ meters per second, or about 40251 kilometers per hour. This speed is called the escape velocity. Notice that the mass of the object, m, canceled out at the last step; the escape velocity is the same for all objects. Of course, it takes considerably more energy to get a large object up to 40251 kph than a small one, so it is certainly more difficult to get a large object into deep space than a small one. Also, note that while we have computed the escape velocity for the earth, this speed would not in fact get an object "to infinity" because of the large mass in our neighborhood called the sun. Escape velocity for the sun starting at the distance of the earth from the sun is nearly 4 times the escape velocity we have calculated.}

Exercises 11.2.
1. Is the area under \( y = 1/x \) from 1 to infinity finite or infinite? If finite, compute the area. \( \Rightarrow \)
2. Is the area under \( y = 1/x^2 \) from 1 to infinity finite or infinite? If finite, compute the area. \( \Rightarrow \)
3. Does \( \int_0^1 x^2 \; dx \) converge or diverge? If it converges, find the value. \( \Rightarrow \)
4. Does \( \int_0^1 \sqrt{x} \; dx \) converge or diverge? If it converges, find the value. \( \Rightarrow \)
5. Does \( \int_0^1 x^2 \; dx \) converge or diverge? If it converges, find the value. \( \Rightarrow \)
6. \( \int_0^\infty \frac{1}{1 + x^2} \; dx \) is an improper integral of a slightly different sort. Express it as a limit and determine whether it converges or diverges; if it converges, find the value. \( \Rightarrow \)
7. Does \( \int_0^\infty \frac{1}{\sqrt{x}} \; dx \) converge or diverge? If it converges, find the value. \( \Rightarrow \)
8. Does \( \int_0^\infty x^3 \; dx \) converge or diverge? If it converges, find the value. \( \Rightarrow \)
9. Does \( \int_0^\infty \frac{1}{x^2 + 1} \; dx \) converge or diverge? If it converges, find the value. \( \Rightarrow \)
10. Does \( \int_0^\infty x \; dx \) converge or diverge? If it converges, find the value. Also find the Cauchy Principal Value, if it exists. \( \Rightarrow \)
11. Does \( \int_0^\infty \sin x \; dx \) converge or diverge? If it converges, find the value. Also find the Cauchy Principal Value, if it exists. \( \Rightarrow \)

11.3 Probability

The study of probability is concerned with more difficult questions as well; for example, suppose the two dice are rolled many times. On the average, what sum will come up? In the language of probability, this average is called the expected value of the sum. This is at first a little misleading, as it does not tell us what to "expect" when the two dice are rolled, but what we expect the long term average will be.

Suppose that two dice are rolled 36 million times. Based on the probabilities, we would expect about 1 million rolls to be 2, about 2 million to be 3, and so on, with a roll of 7 topping the list at about 6 million. The sum of all rolls would be 1 million times 2 plus 2 million times 3, and so on, and dividing by 36 million we would get the average:

\[
x = (2 \cdot 10^6 + 3 \cdot 2 \cdot 10^5 + \ldots + 7 \cdot 10^3 + \ldots + 12 \cdot 10^1) / 36 \cdot 10^6
\]

\[
= 2 \cdot \frac{10^6}{3} + 3 \cdot \frac{2 \cdot 10^6}{3} + \ldots + 7 \cdot \frac{10^3}{36} + \ldots + 12 \cdot \frac{10^1}{36}
\]

\[
= 2P(2) + 3P(3) + \ldots + 7P(7) + \ldots + 12P(12)
\]

\[
= \frac{1}{12} \sum_{i=2}^{12} P(i) = 7.
\]

There is nothing special about the 36 million in this calculation. No matter what the number of rolls, once we simplify the average, we get the same \( \sum_{i=2}^{12} P(i) \). While the actual average value of a large number of rolls will not be exactly 7, the average should be close to 7 when the number of rolls is large. Turning this around, if the average is not close to 7, we should suspect that the dice are not fair.

A variable, say \( X \), that can take certain values, each with a corresponding probability, is called a random variable. In the example above, the random variable was the sum of the two dice. If the possible values for \( X \) are \( x_1, x_2, \ldots, x_n \), then the expected value of the random variable is \( E(X) = \sum_{i=1}^{n} x_i P(x_i) \). The expected value is also called the mean.

When the number of possible values for \( X \) is finite, we say that \( X \) is a discrete random variable. In many applications of probability, the number of possible values of a random variable is very large, perhaps even infinite. To deal with the infinite case we need a different approach, and since there is a sum involved, it should not be wholly surprising that integration turns out to be a useful tool. It then turns out that even when the number of possible values is large but finite, it is frequently easier to pretend that the number is infinite. Suppose, for example, that a dart is thrown at a dart board. Since the dart board consists of a finite number of atoms, there are in some sense only a finite number of places for the dart to land, but it is easier to explore the probabilities involved by pretending that the dart can land on any point in the usual xy plane.
11.3 Probability

DEFINITION 11.3.1 Let \( f : \mathbb{R} \to \mathbb{R} \) be a function. If \( f(x) \geq 0 \) for every \( x \) and \( \int_{-\infty}^{\infty} f(x) \, dx = 1 \) then \( f \) is a probability density function.

We associate a probability density function with a random variable \( X \) by stipulating that the probability that \( X \) is between \( a \) and \( b \) is \( \int_{a}^{b} f(x) \, dx \). Because of the requirement that the integral from \(-\infty\) to \( \infty \) be 1, all probabilities are less than or equal to 1, and the probability that \( X \) takes on some value between \(-\infty\) and \( \infty \) is 1, as it should be.

EXAMPLE 11.3.2 Consider again the two dice example; we can view it in a way that more resembles the probability density function approach. Consider a random variable \( X \) that takes on any real value with probabilities given by the probability density function in figure 11.3.1. The function \( f \) consists of just the top edges of the rectangles, with vertical sides drawn for clarity; the function is zero below 1.5 and above 12.5. The area of each rectangle is the probability of rolling the sum in the middle of the bottom of the rectangle, or

\[
P(n) = \frac{\text{area of rectangle}}{\text{total area}} = f(x) \, dx.
\]

The probability of rolling a 4, 5, or 6 is

\[
P(4) = \frac{1}{6},
\]

\[
P(5) = \frac{1}{6},
\]

\[
P(6) = \frac{1}{6}.
\]

Of course, we could also compute probabilities that don’t make sense in the context of the dice, such as the probability that \( X \) is between 4 and 5.8.

![Figure 11.3.1 A probability density function for two dice.](image)

11.3 Probability

The bell curve.

![Figure 11.3.2 The bell curve.](image)

giving rise to the standard normal distribution. See figure 11.3.2 for the graph of the bell curve.

We have shown that \( A \) is some finite number without computing it; we cannot compute it with the techniques we have available. By using some techniques from multivariable calculus, it can be shown that \( A = \sqrt{2\pi} \).

EXAMPLE 11.3.5 The exponential distribution has probability density function

\[
f(x) = \begin{cases} \frac{1}{\mu} e^{-x/\mu} & x < 0 \\ 0 & x \geq 0 \end{cases}
\]

where \( \mu \) is a positive constant.

The mean or expected value of a random variable is quite useful, as hinted at in our discussion of dice. Recall that the mean for a discrete random variable is \( E(X) = \sum_{i=1}^{n} x_i \cdot P(x_i) \). In the more general context we use an integral in place of the sum.

DEFINITION 11.3.6 The mean of a random variable \( X \) with probability density function \( f \) is \( \mu = E(X) = \int_{-\infty}^{\infty} x \cdot f(x) \, dx \), provided the integral converges.

When the mean exists it is unique, since it is the result of an explicit calculation. The mean does not always exist.

The mean might look familiar; it is essentially identical to the center of mass of a one-dimensional beam, as discussed in section 11.1. The probability density function \( f \) plays the role of the physical density function, but now the “beam” has infinite length. If we consider only a finite portion of the beam, say between \( a \) and \( b \), then the center of mass is

\[
x = \frac{\int_{a}^{b} x \cdot f(x) \, dx}{\int_{a}^{b} f(x) \, dx}
\]

236 Chapter 11 More Applications of Integration

The function

\[
F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t) \, dt
\]

is called the cumulative distribution function or simply (probability) distribution.

EXAMPLE 11.3.3 Suppose that \( a \leq x \leq b \) and

\[
f(x) = \frac{1}{b - a} \quad \text{if } a \leq x \leq b.
\]

Then \( f(x) \) is the uniform probability density function on \( [a, b] \) and the corresponding distribution is the uniform distribution on \( [a, b] \).

EXAMPLE 11.3.4 Consider the function \( f(x) = e^{-x^2/2} \). What can we say about \( \int_{-\infty}^{\infty} e^{-x^2/2} \, dx \)? We cannot find an antiderivative of \( f \), but we can see that this integral is some finite number. Notice that \( 0 < f(x) = e^{-x^2/2} \leq e^{-x^2} \) for \( |x| > 1 \). This implies that the area under \( e^{-x^2/2} \) is less than the area under \( e^{-x^2} \), over the interval \([1, \infty)\). It is easy to compute the latter area, namely

\[
\int_{1}^{\infty} e^{-x^2/2} \, dx = \frac{1}{\sqrt{2\pi}}
\]

so

\[
\int_{-\infty}^{\infty} e^{-x^2/2} \, dx = 2 \int_{0}^{\infty} e^{-x^2} \, dx.
\]

This means that

\[
\int_{-\infty}^{\infty} e^{-x^2/2} \, dx = \int_{-\infty}^{0} e^{-x^2/2} \, dx + \int_{0}^{1} e^{-x^2/2} \, dx + \int_{1}^{\infty} e^{-x^2/2} \, dx = A
\]

for some finite positive number \( A \). Note if we let \( g(x) = f(x)/A \),

\[
\int_{-\infty}^{\infty} g(x) \, dx = \frac{1}{A} \int_{-\infty}^{\infty} e^{-x^2/2} \, dx = \frac{1}{A} A = 1,
\]

so \( g \) is a probability density function. It turns out to be very useful, and is called the standard normal probability density function or more informally the bell curve.

11.3 Probability

If we extend the beam to infinity, we get

\[
x = \int_{-\infty}^{\infty} x \cdot f(x) \, dx = \int_{-\infty}^{\infty} x \cdot f(x) \, dx = E(X),
\]

because \( \int_{-\infty}^{\infty} f(x) \, dx = 1 \). In the center of mass interpretation, this integral is the total mass of the beam, which is always 1 when \( f \) is a probability density function.

EXAMPLE 11.3.7 The mean of the standard normal distribution is

\[
\int_{-\infty}^{\infty} x \cdot e^{-x^2/2} \, dx.
\]

We compute the two halves:

\[
\int_{0}^{\infty} e^{-x^2/2} \, dx = \lim_{b \to \infty} \int_{0}^{b} e^{-x^2/2} \, dx = \frac{1}{\sqrt{2\pi}}
\]

and

\[
\int_{-\infty}^{0} e^{-x^2/2} \, dx = \lim_{b \to -\infty} \int_{b}^{0} e^{-x^2/2} \, dx = \frac{1}{\sqrt{2\pi}}.
\]

The sum of these is 0, which is the mean.

While the mean is very useful, it typically is not enough information to properly evaluate a situation. For example, suppose we could manufacture an 11-sided die, with the faces numbered 2 through 12 so that each face is equally likely to be down when the die is rolled. The value of a roll is the value on this lower face. Rolling the die gives the same range of values as rolling two ordinary dice, but now each value occurs with probability 1/11. The expected value of a roll is

\[
\frac{2}{11} + \frac{3}{11} + \cdots + \frac{12}{11} = 7.
\]

The mean does not distinguish the two cases, though of course they are quite different.

If \( f \) is a probability density function for a random variable \( X \), with mean \( \mu \), we would like to measure how far a “typical” value of \( X \) is from \( \mu \). One way to measure this distance
is \((X - \mu)^2\), we square the difference so as to measure all distances as positive. To get the typical such squared distance, we compute the mean. For two dice, for example, we get
\[
(2 - 7)^2 = \frac{1}{11} + (3 - 7)^2 = \frac{4}{11} + \cdots + (11 - 7)^2 = \frac{6}{11} + (12 - 7)^2 = \frac{35}{11}.
\]
Because we squared the differences this does not measure the typical distance we seek: if we take the square root of this we do get such a measure, \(\sqrt{\text{Var}(X)}\).

To compute the antiderivative, use integration by parts, with \(u = x^2\) and \(dv = e^{-x^2/2} dx\). This gives
\[
\int x e^{-x^2/2} dx = -x e^{-x^2/2} + \int e^{-x^2/2} dx.
\]
We cannot do the new integral, but we know its value when the limits are \(-\infty\) and \(\infty\), namely this one:
\[
\int_{-\infty}^{\infty} x e^{-x^2/2} dx = -\left[\frac{-x e^{-x^2/2}}{2} + \int e^{-x^2/2} dx\right]_{-\infty}^{\infty} = 0 + \frac{1}{\sqrt{2\pi}} \approx 0.242.
\]
Thus the antiderivative is \(\frac{1}{\sqrt{2\pi}} e^{-x^2/2}\).

We cannot do the new integral, but we know its value when the limits are \(-\infty\) and \(\infty\), namely this one:
\[
\int_{-\infty}^{\infty} x e^{-x^2/2} dx = -\left[\frac{-x e^{-x^2/2}}{2} + \int e^{-x^2/2} dx\right]_{-\infty}^{\infty} = 0 + \frac{1}{\sqrt{2\pi}} \approx 0.242.
\]
Thus the antiderivative is \(\frac{1}{\sqrt{2\pi}} e^{-x^2/2}\).

EXAMPLE 11.3.8 We compute the standard deviation of the standard normal distribution. Thus
\[
\int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx = 2\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}.
\]
So there is an 11% chance that this happens—not large, but not tiny. Hence the 15 defective chips example.

How do we measure how unlikely it is that under normal circumstances we would see 15 defective chips? We can’t compute the probability of exactly 15 defective chips, as this would be \(\int_{15}^{15} f(x) dx = 0\). We cannot compute \(\int_{14.5}^{15.5} f(x) dx \approx 0.036\), so if the number of defective chips is 15. (We cannot compute these integrals exactly; computer software has been used to approximate the integral values in this discussion.) But this is misleading: \(\int_{15}^{15.5} f(x) dx \approx 0.126\), which is larger, certainly, but still small, even for the “most likely” outcome. The most useful question, in most circumstances, is this: how likely is it that the number of defective chips is “far from” the mean? For example, how likely, or unlikely, is it that the number of defective chips is different by 5 or more from the expected value of 10? This is the probability that the number of defective chips is less than 5 or larger than 15, namely
\[
\int_{-\infty}^{10} f(x) dx + \int_{15}^{\infty} f(x) dx \approx 0.11.
\]
So there is an 11% chance that this happens—not large, but not tiny. Hence the 15 defective chips do not appear to cause alarm: about one in nine we would expect to see the number of defective chips 5 or more away from the expected 10. How about 20? Here we compute
\[
\int_{-\infty}^{\infty} f(x) dx + \int_{20}^{\infty} f(x) dx \approx 0.0055.
\]
So there is only a 0.15% chance that the number of defective chips is more than 10 away from the mean; this would typically be interpreted as too suspicious to ignore—it shouldn’t happen if the process is running normally.

The big question, of course, is what level of improbability should trigger concern? It depends to some degree on the application, and in particular on the consequences of getting it wrong in one direction or the other. If we’re wrong, do we lose a little money? A lot of money? Do people die? In general, the standard choices are 5% and 1%. So what we should do is find the number of defective chips that has, let us say, a 1% chance of occurring under normal circumstances, and use that as the relevant number. In other words, we want to know when
\[
\int_{-\infty}^{0} f(x) dx + \int_{15}^{\infty} f(x) dx < 0.01.
\]
A bit of trial and error shows that with \(\sigma = 2\) the value is about 0.011, and with \(\sigma = 9\) it is about 0.004, so if the number of defective chips is 19 or more, or 1 or fewer, we should look for problems. If the number is high, we worry that the manufacturing process has a problem, or conversely that the process that tests for defective chips is not working correctly and is flagging good chips as defective. If the number is too low, we suspect that the testing procedure is broken, and is not detecting defective chips.

Exercises 11.3.
1. Verify that \(\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{\pi}\).
2. Show that the function in example 11.3.5 is a probability density function. Compute the mean and standard deviation. \(\Rightarrow\)
3. Compute the mean and standard deviation of the uniform distribution on \([a, b]\). (See example 11.3.5.) \(\Rightarrow\)
4. What is the expected value of one roll of a fair six-sided die? \(\Rightarrow\)
5. What is the expected sum of one roll of three fair six-sided dice? \(\Rightarrow\)
6. Let \(\mu\) and \(\sigma\) be real numbers with \(\sigma > 0\). Show that
\[
N(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x - \mu)^2}{2\sigma^2}}
\]
is a probability density function. You will not be able to compute this integral directly, use a substitution to convert the integral into the one from example 11.3.4. The function \(N\)

Figure 11.3.3 Normal density function for the defective chips example.

Now how do we measure how unlikely it is that under normal circumstances we would see 15 defective chips? We can’t compute the probability of exactly 15 defective chips, as this would be \(\int_{15}^{15} f(x) dx = 0\). We cannot compute \(\int_{14.5}^{15.5} f(x) dx \approx 0.036\), so if the number of defective chips is 15. (We cannot compute these integrals exactly; computer software has been used to approximate the integral values in this discussion.) But this is misleading: \(\int_{15}^{15.5} f(x) dx \approx 0.126\), which is larger, certainly, but still small, even for the “most likely” outcome. The most useful question, in most circumstances, is this: how likely is it that the number of defective chips is “far from” the mean? For example, how likely, or unlikely, is it that the number of defective chips is different by 5 or more from the expected value of 10? This is the probability that the number of defective chips is less than 5 or larger than 15, namely
\[
\int_{-\infty}^{10} f(x) dx + \int_{15}^{\infty} f(x) dx \approx 0.11.
\]
So there is an 11% chance that this happens—not large, but not tiny. Hence the 15 defective chips do not appear to cause alarm: about one in nine we would expect to see the number of defective chips 5 or more away from the expected 10. How about 20? Here we compute
\[
\int_{-\infty}^{\infty} f(x) dx + \int_{20}^{\infty} f(x) dx \approx 0.0055.
\]
So there is only a 0.15% chance that the number of defective chips is more than 10 away from the mean; this would typically be interpreted as too suspicious to ignore—it shouldn’t happen if the process is running normally.

The big question, of course, is what level of improbability should trigger concern? It depends to some degree on the application, and in particular on the consequences of getting it wrong in one direction or the other. If we’re wrong, do we lose a little money? A lot of money? Do people die? In general, the standard choices are 5% and 1%. So what we should do is find the number of defective chips that has, let us say, a 1% chance of occurring under normal circumstances, and use that as the relevant number. In other words, we want to know when
\[
\int_{-\infty}^{0} f(x) dx + \int_{15}^{\infty} f(x) dx < 0.01.
\]
A bit of trial and error shows that with \(\sigma = 2\) the value is about 0.011, and with \(\sigma = 9\) it is about 0.004, so if the number of defective chips is 19 or more, or 1 or fewer, we should look for problems. If the number is high, we worry that the manufacturing process has a problem, or conversely that the process that tests for defective chips is not working correctly and is flagging good chips as defective. If the number is too low, we suspect that the testing procedure is broken, and is not detecting defective chips.

Exercises 11.3.
1. Verify that \(\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{\pi}\).
2. Show that the function in example 11.3.5 is a probability density function. Compute the mean and standard deviation. \(\Rightarrow\)
3. Compute the mean and standard deviation of the uniform distribution on \([a, b]\). (See example 11.3.5.) \(\Rightarrow\)
4. What is the expected value of one roll of a fair six-sided die? \(\Rightarrow\)
5. What is the expected sum of one roll of three fair six-sided dice? \(\Rightarrow\)
6. Let \(\mu\) and \(\sigma\) be real numbers with \(\sigma > 0\). Show that
\[
N(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x - \mu)^2}{2\sigma^2}}
\]
is a probability density function. You will not be able to compute this integral directly, use a substitution to convert the integral into the one from example 11.3.4. The function \(N\)
Now if the graph of $f$ is “nice” (say, differentiable) it appears that we can approximate the length of a portion of the curve with line segments, and that as the number of segments increases, and their lengths decrease, the sum of the lengths of the line segments will approach the true arc length; see figure 11.4.2.

Now we need to write a formula for the sum of the lengths of the line segments, and that as the number of segments increases, and their lengths decrease, the sum of the lengths of the line segments will approach the true arc length; see figure 11.4.2.

As usual, the question is: how might we approximate the surface area? For a surface obtained by rotating a curve around an axis, we can take a polygonal approximation to the curve, as in the last section, and rotate it around the same axis. This gives a surface composed of many “truncated cones;” a truncated cone is called a frustum of a cone. Figure 11.5.1 illustrates this approximation.

Figure 11.5.1  Approximating a surface (left) by portions of cones (right).

So we need to able to compute the area of a frustum of a cone. Since the frustum can be formed by removing a small cone from the top of a larger one, we can compute the desired area if we know the surface area of a cone. Suppose a right circular cone has base radius $r$ and slant height $h$. If we cut the cone from the vertex to the base circle and flatten it out, we obtain a sector of a circle with radius $h$ and arc length $2\pi r$, as in figure 11.5.2. The angle at the center, in radians, is then $2\pi r/h$, and the area of the cone is equal to the area of the sector of the circle. Let $A$ be the area of the sector; since the area of the entire circle is $\pi h^2$, we have

$$A = \frac{2\pi r}{2\pi} \times \frac{2\pi r}{2\pi} h = \pi r h.$$

Now suppose we have a frustum of a cone with slant height $h$ and radii $r_0$ and $r_1$, as in figure 11.5.3. The area of the entire cone is $\pi r_1 h_0$, and the area of the small cone is $\pi r_0 h_0$, thus, the area of the frustum is $\pi r_1 h_0 - \pi r_0 h_0 = \pi ((r_1 - r_0)h_0 + r_1 h_0).$

Figure 11.5.2  The area of a cone.

similar triangles,

$$\frac{h_0}{r_1} = \frac{h + h_1}{r_0},$$

With a bit of algebra this becomes $(r_1 - r_0)h_0 = r_0 h_1.$ Substitution into the area gives

$$\pi((r_1 - r_0)h_0 + r_1 h_1) = \pi(r_0 h_1 + r_1 h) = \pi h(r_0 + r_1) = 2\pi \frac{r_0 + r_1}{2} h = 2\pi h.$$
curve is rotated around the $y$-axis, it forms a frustum of a cone. The area is

$$2πh = 2π\left(\frac{x_i + x_{i+1}}{2}\right)\sqrt{1 + (f'(x))^2} \Delta x.$$

Here $\sqrt{1 + (f'(x))^2} \Delta x$ is the length of the line segment, as we found in the previous section. Assuming $f$ is a continuous function, there must be some $x_i$ in $[x_i, x_{i+1}]$ such that $(f(x_i) + f(x_{i+1}))/2 = f(x_i)$, so the approximation for the surface area is

$$\sum_{i=0}^{n-1} 2πf(x_i)\sqrt{1 + (f'(x))^2} \Delta x.$$

This is not quite the sort of sum we have seen before, as it contains two different values in the interval $[x_i, x_{i+1}]$, namely $x_i$ and $t_i$. Nevertheless, using more advanced techniques than we have available here, it turns out that

$$\lim_{n\to\infty} \sum_{i=0}^{n-1} 2πf(x_i)\sqrt{1 + (f'(x))^2} \Delta x = \int_a^b 2πf(x)\sqrt{1 + (f'(x))^2} \, dx$$

is the surface area we seek. (Roughly speaking, this is because while $x_i$ and $t_i$ are distinct values in $[x_i, x_{i+1}]$, they get closer and closer to each other as the length of the interval shrinks.)

**Figure 11.5.4** One subinterval.

**EXAMPLE 11.5.1** We compute the surface area of a sphere of radius $r$. The sphere can be obtained by rotating the graph of $f(x) = \sqrt{r^2 - x^2}$ about the $x$-axis. The derivative

$$f'(x) = -x/\sqrt{r^2 - x^2},$$

so the surface area is given by

$$A = 2π\int_a^b \sqrt{r^2 - x^2} \sqrt{1 + \left(\frac{-x}{\sqrt{r^2 - x^2}}\right)^2} \, dx.$$

If the curve is rotated around the $y$-axis, the formula is nearly identical, because the length of the line segment we use to approximate a portion of the curve doesn’t change. Instead of the radius $f(x)$, we use the new radius $\bar{r} = (x + x_{i+1})/2$, and the surface area integral becomes

$$\int_a^b 2π\bar{r}\sqrt{1 + (f'(x))^2} \, dx.$$

**EXAMPLE 11.5.2** Compute the area of the surface formed when $f(x) = x^2$ between 0 and 2 is rotated around the $y$-axis.

We compute $f'(x) = 2x$, and then

$$2π\int_0^2 x\sqrt{1 + 4x^2} \, dx = \frac{π}{6}(17^{3/2} - 1),$$

by a simple substitution.

**Exercises 11.5.**

1. Compute the area of the surface formed when $f(x) = 2x\sqrt{x^2 - 1}$ between $-1$ and 0 is rotated around the $x$-axis.$\Rightarrow$
2. Compute the surface area of example 11.5.2 by rotating $f(x) = \sqrt{x}$ around the $x$-axis.$\Rightarrow$
3. Compute the area of the surface formed when $f(x) = x^2$ between 1 and 3 is rotated around the $x$-axis.$\Rightarrow$
4. Compute the area of the surface formed when $f(x) = 2 + \cosh(x)$ between 0 and 1 is rotated around the $x$-axis.$\Rightarrow$
5. Consider the surface obtained by rotating the graph of $f(x) = 1/x$, $x \geq 1$, around the $x$-axis. This surface is called Gabriel’s horn or Torricelli’s trumpet. In exercise 13 in section 11.2 we saw that Gabriel’s horn has finite volume. Show that Gabriel’s horn has infinite surface area.$\Rightarrow$
6. Consider the circle $(x - 2)^2 + y^2 = 1$. Sketch the surface obtained by rotating this circle about the $y$-axis. (The surface is called a torus.) What is the surface area?$\Rightarrow$