11.1 How to prove the formulae for sums of squares and cubes

Mathematicians are concerned with rigorously establishing formulae such as sums of squared (or cubed) integers. While it is not hard to see that these formulae “work” for a few cases, determining that they work in general requires more work. Here we provide a taste of how such careful arguments works. We give two examples. The first, based on mathematical induction provides a general method that could be used in many similar kinds of proofs. The second argument, also for purposes of illustration uses a “trick”. Devising such tricks is not as straightforward, and depends to some degree on serendipity or experience with numbers.

Proof by induction (optional)

Here, we prove the formula for the sum of square integers,

$$\sum_{k=1}^{N} k^2 = \frac{N(N+1)(2N+1)}{6},$$

using a technique called induction. The idea of the method is to check that the formula works for one or two simple cases (e.g. the “sum” of just one or just two terms of the series), and then show that whenever it works for one case (summing up to $N$), it has to also work for the next case (summing up to $N + 1$).

First, we verify that this formula works for a few test cases:

$N = 1$: If there is only one term, then clearly, by inspection,

$$\sum_{k=1}^{1} k^2 = 1^2 = 1.$$
The formula indicates that we should get
\[
S = \frac{1(1 + 1)(2 \cdot 1 + 1)}{6} = \frac{1(2)(3)}{6} = 1,
\]
so this case agrees with the prediction.

\[N = 2;\]
\[
\sum_{k=1}^{2} k^2 = 1^2 + 2^2 = 1 + 4 = 5.
\]
The formula would then predict that
\[
S = \frac{2(2 + 1)(2 \cdot 2 + 1)}{6} = \frac{2(3)(5)}{6} = 5.
\]
So far, elementary computation matches with the result predicted by the formula.

Now we show that if this formula holds for any one case, e.g. for the sum of the first \(N\) squares, then it is also true for the next case, i.e. for the sum of \(N + 1\) squares. So we will assume that someone has checked that for some particular value of \(N\) it is true that
\[
S_N = \sum_{k=1}^{N} k^2 = \frac{N(N + 1)(2N + 1)}{6}.
\]
Now the sum of the first \(N + 1\) squares will be just a bit bigger: it will have one more term added to it:
\[
S_{N+1} = \sum_{k=1}^{N+1} k^2 = \sum_{k=1}^{N} k^2 + (N + 1)^2.
\]
Thus
\[
S_{N+1} = \frac{N(N + 1)(2N + 1)}{6} + (N + 1)^2.
\]
Combining terms, we get
\[
S_{N+1} = (N + 1) \left[ \frac{N(2N + 1)}{6} + (N + 1) \right],
\]
\[
S_{N+1} = (N + 1) \left[ \frac{2N^2 + N + 6N + 6}{6} \right] = (N + 1) \frac{2N^2 + 7N + 6}{6}.
\]
Simplifying and factoring the last term leads to
\[
S_{N+1} = (N + 1) \frac{(2N + 3)(N + 2)}{6}.
\]
We want to check that this still agrees with what the formula predicts. To make the notation simpler, we will let \(M\) stand for \(N + 1\). Then, expressing the result in terms of the quantity \(M = N + 1\) we get
\[
S_M = \sum_{k=1}^{M} k^2 = (N + 1) \frac{2(N + 1) + 1)[(N + 1) + 1]}{6} = M \frac{2M + 1)(M + 1)}{6}.
\]
This is the same formula as we started with, only written in terms of \(M\) instead of \(N\). Thus we have verified that the formula works. By Mathematical Induction we find that the result has been proved.
Another method using a trick

There is another method for determining the sums \( \sum_{k=1}^{n} k^2 \) or \( \sum_{k=1}^{n} k^3 \). Write

\[(k + 1)^3 - (k - 1)^3 = 6k^2 + 2,\]

so

\[\sum_{k=1}^{n} ((k + 1)^3 - (k - 1)^3) = \sum_{k=0}^{n} (6k^2 + 2).\]

But looking more carefully at the left hand side (LHS), we see that

\[\sum_{k=1}^{n} ((k + 1)^3 - (k - 1)^3) = 2^3 - 0^3 + 3^3 - 1^3 + 4^3 - 2^3 + 5^3 - 3^3 \ldots + (n + 1)^3 - (n - 1)^3.\]

most of the terms cancel, leaving only \(-1 + n^3 + (n + 1)^3\), so this means that

\[-1 + n^3 + (n + 1)^3 = 6 \sum_{k=1}^{n} k^2 + \sum_{k=1}^{n} 2,\]

so

\[\sum_{k=1}^{n} k^2 = (-1 + n^3 + (n + 1)^3 - 2n)/6 = (2n^3 + 3n^2 + n)/6.\]

Similarly, the formula for \( \sum_{k=1}^{n} k^3 \), can be obtained by starting with

\[(k + 1)^4 - (k - 1)^4 = 4k^3 + 4k.\]

11.2 Riemann Sums: Extensions and other examples

We take up some issues here that were not yet considered in the context of our examples of Riemann sums in Chapter 2. First, we consider an arbitrary interval \( a \leq x \leq b \). Then we comment on other ways of constructing the rectangular strip approximation (that eventually lead to the same limit when the true area is computed.)

11.2.1 A general interval: \( a \leq x \leq b \)

Example 2: (Lu Fan)

Find the area under the graph of the function

\[y = f(x) = x^2 + 2x + 1 \quad a \leq x \leq b.\]

\(\text{I want to thank Robert Israel for contributing this material}\)
Here the interval is \( a \leq x \leq b \). Let us leave the values of \( a, b \) general for a moment, and consider how the calculation is set up in this case. Then we have

- length of interval = \( b - a \)
- number of segments = \( N \)
- width of rectangular strips = \( \Delta x = \frac{b - a}{N} \)
- the \( k \)’th \( x \) value = \( x_k = a + k \frac{(b - a)}{N} \)
- height of \( k \)’th rectangular strip = \( f(x_k) = x_k^2 + 2x_k + 1 \)

Combining the last two steps, the height of rectangle \( k \) is:

\[
f(x_k) = \left( a + \frac{k(b - a)}{N} \right)^2 + 2 \left( a + \frac{k(b - a)}{N} \right) + 1
\]

and its area is

\[
a_k = f(x_k) \times \Delta x = f(x_k) \times \left( \frac{b - a}{N} \right).
\]

We use the last two equations to express \( a_k \) in terms of \( k \) (and the quantities \( a, b, N \)), then sum over \( k \) as before (\( A = \sum A_k \)). Some algebra is needed to simplify the sums so that summation formulae can be applied. The details are left as an exercise for the reader (see homework problems). Evaluating the limit \( N \to \infty \), we finally obtain

\[
A = \lim_{N \to \infty} \sum_{k=1}^{N} a_k = (a + 1)^2(b - a) + (a + 1)(b - a)^2 + \frac{(b - a)^3}{3}.
\]

as the area under the function \( f(x) = x^2 + 2x + 1 \), over the interval \( a \leq x \leq b \). Observe that the solution depends on \( a \), and \( b \). (The endpoints of the interval influence the total area under the curve.) For example, if the given interval happens to be \( 2 \leq x \leq 4 \). then, substituting \( a = 2 \), \( b = 4 \) into the above result for \( A \), leads to

\[
A = (2 + 1)^2(4 - 2) + (2 + 1)(4 - 2)^2 + \frac{4 - 2}{3} = 18 + 12 + \frac{2}{3} = \frac{32}{3}
\]

In the next chapter, we will show that the tools of integration will lead to the same conclusion.

### 11.2.2 Using left (rather than right) endpoints

So far, we used the right endpoint of each rectangular strip to assign its height using the given function (see Figs. 2.2, 2.3, 2.4). Restated, we “glued” the top right corner of the rectangle to the graph of the function. This is the so called right endpoint approximation. We can just as well use the left corners of the rectangles to assign their heights (left endpoint approximation). A comparison of these for the function \( y = f(x) = x^2 \) is shown in Figs. 11.1 and 11.2. In the case of the left endpoint approximation, we evaluate
Figure 11.1. The area under the curve $y = f(x)$ over an interval $a \leq x \leq b$ could be computed by using either a left or right endpoint approximation. That is, the heights of the rectangles are adjusted to match the function of interest either on the right or on their left corner. Here we compare the two approaches. Usually both lead to the same result once a limit is computer to arrive at the “true” area.

the heights of the rectangles starting at $x_0$ (instead of $x_1$, and ending at $x_{N-1}$ (instead of $x_N$). There are still $N$ rectangles. To compare, sum of areas of the rectangles in the left versus the right endpoint approximation is

\[
\text{Right endpoints: } A_N \text{ strips } = \sum_{k=1}^{N} f(x_k) \Delta x.
\]

\[
\text{Left endpoints: } A_N \text{ strips } = \sum_{k=0}^{N-1} f(x_k) \Delta x.
\]

Details of one such computation is given in the box.
Example of left endpoint calculation

We here look again at a simple example, using the quadratic function,

\[ f(x) = x^2, \quad 0 \leq x \leq 1, \]

We now compare the right and left endpoint approximation. These are shown in panels of Figure 11.2. Note that

\[ \Delta x = \frac{1}{N}, \quad x_k = \frac{k}{N}, \]

The area of the \( k \)’th rectangle is

\[ a_k = f(x_k) \times \Delta x = \left( \frac{k}{N} \right)^2 \left( \frac{1}{N} \right), \]

but now the sum starts at \( k = 0 \) so

\[ A_N \text{ strips} = \sum_{k=0}^{N-1} f(x_k) \Delta x = \sum_{k=0}^{N-1} \left( \frac{k}{N} \right)^2 \left( \frac{1}{N} \right) = \left( \frac{1}{N^3} \right) \sum_{k=0}^{N-1} k^2. \]

The first rectangle corresponds to \( k = 0 \) in the left endpoint approximation (rather than \( k = 1 \) in the right endpoint approximation). But the \( k = 0 \) rectangle makes no contribution (as its area is zero in this example) and we have one less rectangle at the right endpoint of the interval, since the \( N \)’th rectangle is \( k = N - 1 \). Then the sum is

\[ A_N \text{ strips} = \left( \frac{1}{N^3} \right) \frac{(2(N-1) + 1)(N-1)(N)}{6} = \frac{(2N - 1)(N - 1)}{6N^2}. \]

The area, obtained by taking a limit for \( N \to \infty \) is

\[ A = \lim_{N \to \infty} A_N \text{ strips} = \lim_{N \to \infty} \frac{(2N - 1)(N - 1)}{6N^2} = \frac{2}{6} = \frac{1}{3}. \]

We see that, after computing the limit, the result for the “true area” under the curve is exactly the same as we found earlier in this chapter using the right endpoint approximation.

11.3 Physical interpretation of the center of mass

We defined the idea of a center of mass in Chapter 5. The center of mass has a physical interpretation for a real mass distribution. Loosely speaking, it is the position at which the mass “balances” without rotating to the left or right. In physics, we say that there is no net torque. The analogy with children sitting on a teeter-totter is relevant: many children may sit along the length of the frame of a teeter totter, but if they distribute themselves in a way that the center of mass is at the fulcrum of the teeter totter, they will remain precariously balanced (until one of them fidgets or gets off!). Notice that both the mass and the position of each child is important - a light child sitting on the very edge of the teeter totter can balance a heavier child sitting closer to the fulcrum (middle). The center of mass need not be the same as the median position. As we have see, the median is a position that
subdivides the distribution into two equal masses (or, more generally, produces equal sized areas under the graph of the density function.) The center of mass assigns a greater weight to parts of the distribution that are “far away” in the same sense. (However, for symmetric distributions, the median and the mean are the same.)

In physics, we speak of the “moment of mass” of a distribution about a point. This quantity is related to the tendency of the mass to contribute a torque, i.e. to make the object rotate. Suppose we are interested in a particular point of reference $x$. In a discrete mass distribution, for example, the moment of mass of each of the beads relative to point $x$ is given by the product of the mass and its distance away from the point - as with the teeter totter, beads farther away will contribute more torque than beads closer to point $x$, and heavier beads (i.e. greater mass) will contribute more torque than lighter beads. For example, mass 1 contributes an amount $m_1(x - x_1)$ to the total moment of mass of the distribution about the point $x$. Altogether the moment of mass of the distribution about the point $x$ is given by the sum of the contributions from all the beads:

$$M(x) = \sum_{i=1}^{N} m_i(x - x_i).$$

The center of mass is the point about which the moment of mass is zero, i.e.

$$M(x) = 0.$$

This point is given by the formula:

$$x_{cm} = \frac{\sum_{i=1}^{N} m_i x_i}{\sum_{i=1}^{N} m_i}.$$
point \( x \) is defined as
\[
M_1(x) = \sum_{i=1}^{n} m_i (x - x_i).
\]

The center of mass is a special point \( \bar{x} \) such that the moment of mass about that point is zero. (Loosely speaking the tendency to rotate to the left or the right are the same: thus the distribution would be balanced if it “rested on that point”.)

![Diagram of three masses and their center of mass](image)

**Figure 11.3.** A discrete set of masses \( m_1, m_2, m_3 \) is distributed at positions \( x_1, x_2, x_3 \). The center of mass of the distribution is the position at which the given mass distribution would balance, here represented by the white triangle.

Thus, we identify the center of mass as the point at which
\[
M_1(\bar{x}) = 0,
\]
or
\[
\sum_{i=1}^{n} m_i (\bar{x} - x_i) = 0.
\]

Now expanding the sum, we rewrite the above as
\[
\left( \sum_{i=1}^{n} m_i \bar{x} \right) - \left( \sum_{i=1}^{n} m_i x_i \right) = 0,
\]
\[
\bar{x} \sum_{i=1}^{n} m_i - \left( \sum_{i=1}^{n} m_i x_i \right) = 0.
\]

But we already know that the first summation above is just the total mass, so that
\[
\bar{x} M - \left( \sum_{i=1}^{n} m_i x_i \right) = 0,
\]
so, taking the second term to the other side and dividing by \( M \) leads to
\[
\bar{x} = \frac{1}{M} \sum_{i=1}^{n} m_i x_i.
\]

We have recovered precisely the definition of the center of mass or “average \( x \) coordinate”.
11.4 The shell method for computing volumes

In Chapter 5, we used dissection into small disks to compute the volume of solids of revolution. Here we show use an alternative dissection into shells.

11.4.1 Example: Volume of a cone using the shell method

We use the shell method to find the volume of the cone formed by rotating the curve \( y = 1 - x \) about the \( y \) axis.

**Solution**

We show the cone and its generating curve in Figure 11.4, together with a representative shell used in the calculation of total volume. The volume of a cylindrical shell of radius \( r \), height \( h \) and thickness \( \tau \) is

\[
V_{\text{shell}} = 2\pi rh\tau.
\]

We will place these shells one inside the other so that their radii are parallel to the \( x \) axis (so \( r = x \)). The heights of the shells are determined by their \( y \) value (i.e. \( h = y = 1 - x \)).

---

Note to the instructor: This material may be skipped in the interest of time. It presents an alternative to the disk method, but there may not be enough time to cover this in detail.
1 - r). For the tallest shell \( r = 0 \), and for the flattest shell \( r = 1 \). The thickness of the shell is \( \Delta r \). Therefore, the volume of one shell is

\[
V_{\text{shell}} = 2\pi r (1 - r) \Delta r.
\]

The volume of the object is obtained by summing up these shell volumes. In the limit, as \( \Delta r \to dr \) gets infinitesimally small, we recognize this as a process of integration. We integrate over \( 0 \leq r \leq 1 \), to obtain:

\[
V = 2\pi \int_0^1 r (1 - r) \, dr = 2\pi \int_0^1 (r - r^2) \, dr.
\]

We find that

\[
V = 2\pi \left( \frac{r^2}{2} - \frac{r^3}{3} \right) \bigg|_0^1 = 2\pi \left( \frac{1}{2} - \frac{1}{3} \right) = \pi/3.\]

### 11.5 More techniques of integration

#### 11.5.1 Secants and other “hard integrals”

In a previous section, we encountered the integral

\[
I = \int \sec^3(x) \, dx.
\]

This integral can be simplified to some extent by integration by parts as follows: Let \( u = \sec(x), \, dv = \sec^2(x) \, dx \). Then \( du = \sec(x) \tan(x) \, dx \) while \( v = \int \sec^2(x) \, dx = \tan(x) \).

The integral can be transformed to

\[
I = \sec(x) \tan(x) - \int \sec(x) \tan^2(x) \, dx.
\]

The latter can be rewritten:

\[
I_1 = \int \sec(x) \tan^2(x) \, dx = \int \sec(x)(\sec^2(x) - 1).\]

where we have use a trigonometric identity for \( \tan^2(x) \). Then

\[
I = \sec(x) \tan(x) - \int \sec^3(x) \, dx + \int \sec(x) \, dx = \sec(x) \tan(x) - I + \int \sec(x) \, dx
\]

so (taking both I’s to the left hand side, and dividing by 2)

\[
I = \frac{1}{2} \left( \sec(x) \tan(x) + \int \sec(x) \, dx \right).
\]

We are now in need of an antiderivative for \( \sec(x) \). No “obvious substitution” or further integration by parts helps here, but it can be checked by differentiation that

\[
\int \sec(x) \, dx = \ln |\sec(x) + \tan(x)| + C
\]

Then the final result is

\[
I = \frac{1}{2} (\sec(x) \tan(x) + \ln |\sec(x) + \tan(x)|) + C
\]
11.5. More techniques of integration

11.5.2 A special case of integration by partial fractions

Evaluate this integral\(^{72}\):

$$\int_1^2 \frac{7x + 4}{6x^2 + 7x + 2} \, dx$$

This integral involves a rational function (that is, a ratio of two polynomials). The denominator is a degree 2 polynomial function that has two roots and that can be factored easily; the numerator is a degree 1 polynomial function. In this case, we can use the following strategy. First, factor the denominator:

$$6x^2 + 7x + 2 = (2x + 1)(3x + 2)$$

Assign \(A\) and \(B\) in the following way:

$$\frac{A}{2x + 1} + \frac{B}{3x + 2} = \frac{7x + 4}{(2x + 1)(3x + 2)} = \frac{7x + 4}{6x^2 + 7x + 2}$$

(Remember, this is how we define \(A\) and \(B\).)

Next, find the common denominator and rewrite it as a single fraction in terms of \(A\) and \(B\).

$$\frac{A}{2x + 1} + \frac{B}{3x + 2} = \frac{3Ax + 2A + 2Bx + B}{(2x + 1)(3x + 2)}$$

Group like terms in the numerator, and note that this has to match the original fraction, so:

$$\frac{3Ax + 2A + 2Bx + B}{(2x + 1)(3x + 2)} = \frac{(3A + 2B)x + (2A + B)}{(2x + 1)(3x + 2)}$$

The above equation should hold true for all \(x\) values; therefore:

$$3A + 2B = 7, \quad 2A + B = 4$$

Solving the system of equations leads to \(A = 1, B = 2\). Using this result, we rewrite the original expression in the form:

$$\frac{7x + 4}{6x^2 + 7x + 2} = \frac{7x + 4}{(2x + 1)(3x + 2)} = \frac{1}{2x + 1} + \frac{2}{3x + 2}$$

Now we are ready to rewrite the integral:

$$I = \int_1^2 \frac{7x + 4}{6x^2 + 7x + 2} \, dx = \int_1^2 \left( \frac{1}{2x + 1} + \frac{2}{3x + 2} \right) \, dx$$

Simplify:

$$I = \int_1^2 \frac{1}{2x + 1} \, dx + 2 \int_1^2 \frac{1}{3x + 2} \, dx$$

Now the integral becomes a simple natural log integral that follows the pattern of Eqn. 6.1. Simplify:

$$I = \frac{1}{2} \ln |2x + 1| \bigg|_1^2 + \frac{2}{3} \ln |3x + 2| \bigg|_1^2$$

\(^{72}\)This section was contributed by Lu Fan
Chapter 11. Appendix

Simplify further:

\[ I = \frac{1}{2}(\ln 5 - \ln 3) + \frac{2}{3}(\ln 8 - \ln 5) = -\frac{1}{6}\ln 5 - \frac{1}{2}\ln 3 + \frac{2}{3}\ln 8. \]

This method can be used to solve any integral that contain a fraction with a degree 1 polynomial in the numerator and a degree 2 polynomial (that has two roots) in the denominator.

11.6 Analysis of data: a student grade distribution

We study the distribution of student grades on a test written by 76 students and graded out of a maximum of 50 points.

11.6.1 Defining an average grade

Let \( N \) be the size of the class, and \( y_k \) the grade of student \( k \). Here \( k \) is the number of the student from 1 to \( N \), and \( y_k \) takes any value between 0 and 50 points. Then the average grade \( \bar{Y} \) is computed by adding up the scores of all students and dividing by the number of students as follows:

\[ \bar{Y} = \frac{1}{N} \sum_{k=1}^{N} y_k. \]

For example, for a class of 76 students, we would have the sum

\[ \bar{Y} = \frac{1}{76} \sum_{k=1}^{76} y_k. \]

11.6.2 Fraction of students that scored a given grade

Suppose that the number of students who got the grade \( x_i \) is \( p_i \). If the class consists of a total of \( N \) students, then it follows that

\[ N = \sum_{i=1}^{10} p_i. \]

This is just saying that the sum of the number of students in every one of the categories has to add up to the total class size. The fraction of the class that scored grade \( x_i \) is

\[ \frac{p_i}{N}. \]

(Dividing by \( N \) has normalized the distribution. The value \( p_i/N \) is the empirical probability of getting grade \( x_i \).) The mean or average grade is:

\[ \bar{X} = \frac{1}{N} \sum_{i=0}^{50} x_i p_i. \]
11.6. Analysis of data: a student grade distribution

11.6.3 Frequency distribution

It is difficult to visualize all the data if we list all the grades obtained. We “lump together” scores into various categories (or “bins”) and create a distribution. For example, test scores might be divided into ranges of bins in increments of 5 points: (1-5, 6-10, 11-15, etc). We could represent grades in each bin by some value up to a specified level of accuracy. For example, grades in the range 16-20 can be described by the score 18 up to an accuracy of ±2. This is what we have done in Table 11.1.

We will now reinterpret our notation somewhat. We will refer to \( \tilde{x}_i \) as the score and \( p_i \) the number of students whose test score fell within the range represented by \( \tilde{x}_i \pm \) accuracy. (The notation \( \tilde{x}_i \) is meant to remind us that we are approximating the grade value.) For example, consider 10 “bins” or grade categories. In that case, the index \( i \) takes on values \( i = 1, 2, \ldots, 10 \). The, e.g., \( \tilde{x}_4 \) represents all grades in the fourth “bin”, i.e. grades between 16-20. A plot of \( p_i \) against \( \tilde{x}_i \) is called a frequency distribution. The bar graph shown in Figure 11.5 represents this distribution. Table 11.1 shows the data that produced that bar graph.

11.6.4 Average/mean of the distribution

The frequency distribution can also be used to compute an average value: each (approximate) grade value \( \tilde{x}_i \) is achieved by \( p_i \) students, which is a fraction \( (p_i/N) \) of the whole class. When we form the multiple \( (p_i/N)\tilde{x}_i \), we assign a “weight” to each of the cate-
Table 11.1. Distribution of grades (out of 50) for a class of 76 students. The mean grade for this class is 31.9474.

<table>
<thead>
<tr>
<th>i</th>
<th>grade $\tilde{x}_i$</th>
<th>number $p_i$</th>
<th>$\sum p_i$</th>
<th>$\sum \tilde{x}_i p_i$</th>
<th>$(1/N) \sum \tilde{x}_i p_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>1</td>
<td>3±2</td>
<td>1</td>
<td>1</td>
<td>3.0</td>
<td>0.0395</td>
</tr>
<tr>
<td>2</td>
<td>8±2</td>
<td>2</td>
<td>3</td>
<td>19</td>
<td>0.25</td>
</tr>
<tr>
<td>3</td>
<td>13±2</td>
<td>3</td>
<td>3</td>
<td>19</td>
<td>0.25</td>
</tr>
<tr>
<td>4</td>
<td>18±2</td>
<td>5</td>
<td>8</td>
<td>109</td>
<td>1.4342</td>
</tr>
<tr>
<td>5</td>
<td>23±2</td>
<td>18</td>
<td>18</td>
<td>339</td>
<td>4.4605</td>
</tr>
<tr>
<td>6</td>
<td>28±2</td>
<td>8</td>
<td>26</td>
<td>563</td>
<td>7.4079</td>
</tr>
<tr>
<td>7</td>
<td>33±2</td>
<td>21</td>
<td>47</td>
<td>1256</td>
<td>16.5263</td>
</tr>
<tr>
<td>8</td>
<td>38±2</td>
<td>19</td>
<td>66</td>
<td>1978</td>
<td>26.0263</td>
</tr>
<tr>
<td>9</td>
<td>43±2</td>
<td>6</td>
<td>72</td>
<td>2236</td>
<td>29.4211</td>
</tr>
<tr>
<td>10</td>
<td>48±2</td>
<td>4</td>
<td>76</td>
<td>2428</td>
<td>31.9474</td>
</tr>
</tbody>
</table>

We define the mean or average grade in the distribution by

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{M} \tilde{x}_i p_i. \quad (11.1)$$

Where $M$ is the number of bins. An equivalent way of expressing the mean (average) is:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{M} \tilde{x}_i p_i = \frac{\sum_{i=1}^{M} \tilde{x}_i p_i}{\sum_{i=1}^{M} p_i}. \quad (11.2)$$

The sum in the denominator of this last fraction is simply the total class size.

In Table 11.1, we show steps in the calculation of the mean grade for the class. This calculation is easily handled on the same spreadsheet that recorded the frequency of grades and that was used to plot the bar graph of that distribution. Equations 11.1 and 11.2 are saying the same thing. We will see the second of these again in the context of a more general probability distribution in Chapter 8.

### 11.6.5 Cumulative function

We can calculate a “running total” as shown on Figure 11.6, where we plot for each grade category, the total number of students whose grade was in the given range.
We define the cumulative function, $F_i$, to be:

$$ F_i = \sum_{k=1}^{i} p_k. $$

Then $F_i$ is the number of students whose grade $x_k$ was between $x_1$ and $x_i$ ($x_1 \leq x_k \leq x_i$). Of course, when we add up all the way to the last category, we arrive at the total number of students in the class (assuming each student wrote the test and received a grade). Thus

$$ F_m = \sum_{k=1}^{M} p_k = N, $$

Where as before, $M$ stands for the number of “bins” used to represent the grade distribution. (Note that each student has been counted in one of the categories corresponding to the grade he or she achieved.) Another way of saying the same thing is that

$$ \sum_{k=1}^{m} \frac{p_k}{N} = 1. $$

In Figure 11.6 we show the cumulative function, i.e. we plot $\tilde{x}_i$ vs $F_i$. Note that this graph is a step function. That is, the function takes on a set of discrete values with jumps at every 5th integer.

Figure 11.6. Top: The same grade distribution as in Figure 11.5, but showing the cumulative function. The grid has been removed for easier visualization of that step function. Bottom: The cumulative function is used to determine an approximate median grade.

\footnote{Note: ideally, this graph should be discontinuous, with horizontal segments only. The vertical “jumps” cannot correspond to values of a function. However the spreadsheet tool used to plot this function does not currently allow this graphing option.}
11.6.6 The median

We can use the cumulative function and its features to come up with new ways of summarizing the distribution or comparing the performance of two sections. Suppose we subdivide a given class into exactly two equal groups based on performance on the test. Then there would be some grade that was achieved or surpassed by the top half of the class only; the rest of the students (i.e. the other half of the class) got scores below that level. We call that grade the median of the distribution.

To find the median grade using a cumulative function, we must ask what grade level corresponds to a cumulative 1/2 of the class, i.e. to \( N/2 \) students. To determine that level, we draw a horizontal line corresponding to \( N/2 \). As shown in Figure 11.6, because the function \( f \) is discontinuous, we only have an approximate median of 30. We observe that the median is not in general equal to the mean computed earlier.

11.7 Factorial notation

Let \( n \) be an integer, \( n \geq 0 \). Then \( n! \), called “n factorial”, is defined as the following product of integers:

\[
n! = n(n-1)(n-2) \ldots (2)(1)
\]

Example

\[
1! = 1 \\
2! = 2 \cdot 1 = 2 \\
3! = 3 \cdot 2 \cdot 1 = 6 \\
4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24 \\
5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120
\]

We also define

\[
0! = 1
\]

11.8 Appendix: Permutations and combinations

11.8.1 Permutations

A permutation is a way of arranging objects, where the order of appearance of the objects is important.
11.9 Appendix: Tests for convergence of series

In order for the sum of ‘infinitely many things’ to add up to a finite number, the terms have to get smaller. But just getting smaller is not, in itself, enough to guarantee convergence. (We will show this later on by considering the harmonic series.)

There are rigorous mathematical tests which help determine whether a series converges or not. We discuss some of these tests here\footnote{Recall that $\Rightarrow$ means “implies that”. This is a one-way implication: $A \Rightarrow B$ says that “A implies B” and cannot be used to conclude that B implies A. $\iff$ means that each statement implies the other, a two-way implication. Just as it is important to “obey traffic signs” and avoid “driving the wrong way” on a one-way street, it is also important to be careful about use of these mathematical statements.}.

---

\begin{itemize}
\item \textbf{(a)} The number of permutations (ways of arranging) $n$ objects into $n$ slots. There are $n$ choices for the first slot, and for each of these, there are $n-1$ choices for the second slot, etc. In total there are $n!$ ways of arranging these objects. (Note that the order of the objects is here important.)
\item \textbf{(b)} The number of permutations of $n$ objects into $k$ slots, $P(n,k)$, is the product $n \cdot (n-1) \cdot (n-2) \ldots (n-k+1)$ which can also be written as a ratio of factorials.
\item \textbf{(c)} The number of combinations of $n$ objects in groups of $k$ is called $C(n,k)$ (shown as the first arrow in part \textbf{c}). Here order is not important. The step shown in (b) is equivalent to the two steps shown in (c). This means that there is a relationship between $P(n,k)$ and $C(n,k)$, namely, $P(n,k) = k!C(n,k)$.
\end{itemize}
11.9.1 The ratio test:

If \( \sum_{k=0}^{\infty} a_k \) is a series with \( a_n > 0 \) and \( \lim_{k \to \infty} \frac{a_{k+1}}{a_k} = L \), then

(a) \( L < 1 \Rightarrow \) the series converges,
(b) \( L > 1 \Rightarrow \) the series diverges,
(c) \( L = 1 \Rightarrow \) the test is inconclusive.

Example 1: Reciprocal factorial series

Recall that if \( k > 0 \) is an integer then the notation \( k! \) (read “\( k \) factorial”) means

\[ k! = k \cdot (k-1) \cdot (k-2) \ldots 3 \cdot 2 \cdot 1. \]

Consider the series

\[ S = \sum_{k=1}^{\infty} \frac{1}{k!} = 1 + \frac{1}{2 \cdot 1} + \frac{1}{3 \cdot 2 \cdot 1} + \ldots + \frac{1}{k(k-1)\ldots1}, \]

then

\[ a_{k+1} = \frac{1}{(k+1)!}, \quad a_k = \frac{1}{k!}, \]

\[ \frac{a_{k+1}}{a_k} = \lim_{k \to \infty} \frac{1}{(k+1)k!} = \lim_{k \to \infty} \frac{k!}{(k+1)!} = \lim_{k \to \infty} \frac{1}{k+1} = 0. \]

Thus \( L = 0, L < 1 \) so this series converges by the ratio test. Later, we will see a second method (comparison) to arrive at the same conclusion.

Example 2: Harmonic series

Does the following converge?

\[ S = \sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{k} + \ldots, \]

This series is the Harmonic Series. To apply the ratio test, we note that

\[ a_{k+1} = \frac{1}{k+1}, \quad a_k = \frac{1}{k}, \]

\[ \frac{a_{k+1}}{a_k} = \lim_{k \to \infty} \frac{1}{(k+1)k} = \lim_{k \to \infty} \frac{k}{k+1} = 1. \]

Since \( L = 1 \), in this case, the test is inconclusive. In fact, we show in Section 10.4 that the harmonic series diverges.
Example 3: Geometric series

Apply the ratio test to determine the condition for convergence of the geometric series,

\[ S = \sum_{k=0}^{\infty} r^k. \]

Here

\[ a_{k+1} = r^{k+1}, \quad a_k = r^k, \quad \frac{a_{k+1}}{a_k} = r, \]

\[ L = \lim_{k \to \infty} \frac{a_{k+1}}{a_k} = r. \]

So, by the ratio test, if \( L = r < 1 \) then the geometric series converges (confirming a fact we have already established).

11.9.2 Series comparison tests

We can sometimes use the convergence (or divergence) of a known series to conclude whether a second series converges (or diverges).

<table>
<thead>
<tr>
<th>Suppose we have two series,</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_a = \sum_{k=0}^{\infty} a_k ) and ( S_b = \sum_{k=0}^{\infty} b_k, )</td>
</tr>
</tbody>
</table>

such that terms of one series are always smaller than terms of the other, i.e. satisfy

\[ 0 < a_k < b_k \quad \text{for all } k = 0, 1, \ldots. \]

Then

\[ \sum b_k \text{ converges } \Rightarrow \sum a_k \text{ converges}, \]

\[ \sum a_k \text{ diverges } \Rightarrow \sum b_k \text{ diverges}. \]

The idea behind the first of these statements is that the “smaller” series \( \sum a_k \) is “squeezed in” between 0 (the lower bound) and the sum of the larger series (which we know must exist, since \( \sum b_k \) converges.) This means that the smaller series cannot become unbounded. For the second statement, we have that the smaller of the two series is known to diverge, forcing the larger also to be unbounded. One must carefully observe that “\( \Rightarrow \)” applies only in one direction. (For example, if the smaller series converges, we cannot conclude anything about the larger series.)

Example: Comparison with geometric series

Does the series below converge or diverge?

\[ S = \sum_{k=0}^{\infty} \frac{1}{2^k + 1}. \]
Solution: We compare terms in this series to a terms in a geometric series with \( r = \frac{1}{2} \). i.e.
consider
\[
a_k = \frac{1}{2^k + 1}, \quad b_k = \frac{1}{2^k}.
\]
Then clearly
\[
0 < a_k < b_k \quad \text{for every } k
\]
(since the denominator in \( a_k \) is larger). But we know that \( \sum \frac{1}{2^k} \) converges. Therefore, so does \( \sum \frac{1}{2^k + 1} \).

11.9.3 Alternating series

An alternating series is a series in which the signs of successive terms alternate. An example of this type is the series
\[
1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots = \sum (-1)^{n+1} \frac{1}{n}
\]
We will show that this series converges (essentially because terms nearly cancel out), and in fact, we show in Section 10.5.3 that it converges to the number \( \ln(2) \approx 0.693 \). More generally, we have the following result.

If \( S \) is an alternating series,
\[
S = \sum_{k=1}^{\infty} (-1)^k a_k = a_1 - a_2 + a_3 - a_4 + \ldots
\]
with \( a_k > 0 \) and such that (1) \( |a_1| \geq |a_2| \geq |a_3| \geq \ldots \) etc. and (2) \( \lim_{k \to \infty} a_k = 0 \), then
the series converges. (This was established by Leibniz in 1705.)

11.10 Adding and multiplying series

We first comment that arithmetic operations on infinite series only make sense if the series are convergent. In this discussion, we will deal only with series of the convergent type. When this is true, then (and only then) is it true that we can exchange the order of operations as discussed below.

If \( \sum a_k \) and \( \sum b_k \) both converge and \( \sum a_k = S \), \( \sum b_k = T \), then
(a) \( \sum (a_k + b_k) \) converges and \( \sum (a_k + b_k) = \sum a_k + \sum b_k = S + T \).
(b) \( \sum c a_k = c \sum a_k = cS \), where \( c \) is any constant.
(c) The product \( \left( \sum a_k \right) \cdot \left( \sum b_k \right) = \sum_{k=0}^{\infty} \sum_{i=0}^{k} a_i b_{k-i} = S \cdot T \).
Example:

\[
\sum \left( \frac{1}{2} \right)^k \cdot \sum \left( \frac{1}{3} \right)^j = \left( 1 + \frac{1}{2} + \frac{1}{4} + \ldots \right) \left( 1 + \frac{1}{3} + \frac{1}{9} + \ldots \right).
\]

Both series converge, so we can write

\[
\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \left( \frac{1}{2} \right)^k \left( \frac{1}{3} \right)^j = \frac{1}{1 - \frac{1}{2}} \cdot \frac{1}{1 - \frac{1}{3}} = 2 \cdot \frac{3}{2} = 3.
\]

11.11 Using series to solve a differential equation

Airy’s equation arises in the study of optics, and (with initial conditions) is as follows:

\[
y'' = xy, \quad y(0) = 1, \quad y'(0) = 0.
\]

As before, we will write the solution as a series:

\[
y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + a_5 x^5 + \ldots
\]

Using the information from the initial conditions, we get \( y(0) = a_0 = 1 \) and \( y'(0) = a_1 = 0 \). Now we can write down the derivatives:

\[
y' = a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + 5a_5 x^4 + \ldots
\]

\[
y'' = 2a_2 + 2 \cdot 3x + 3 \cdot 4x^2 + 4 \cdot 5x^3 + \ldots
\]

The equation then gives

\[
y'' = xy
\]

\[
2a_2 + 2 \cdot 3a_3 x + 3 \cdot 4a_4 x^2 + 4 \cdot 5a_5 x^3 + \ldots = x(a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots)
\]

Again, we can equate the coefficients of \( x \), and use \( a_0 = 1 \) and \( a_1 = 0 \), to obtain

\[
\begin{align*}
2a_2 &= 0 \quad \Rightarrow a_2 = 0, \\
2 \cdot 3a_3 &= a_0 \quad \Rightarrow a_3 = \frac{1}{3}, \\
3 \cdot 4a_4 &= a_1 \quad \Rightarrow a_4 = 0, \\
4 \cdot 5a_5 &= a_2 \quad \Rightarrow a_5 = 0, \\
5 \cdot 6a_6 &= a_3 \quad \Rightarrow a_6 = \frac{1}{5 \cdot 6}.
\end{align*}
\]

This gives us the first few terms of the solution:

\[
y = 1 + \frac{x^3}{2 \cdot 3} + \frac{x^6}{2 \cdot 3 \cdot 5 \cdot 6} + \ldots
\]

If we continue in this way, we can write down many terms of the series.
Index

3D
  objects, 81

Abel’s theorem, 211
acceleration, 62
actin
cortex, 84
addition
  principle, 140
age
distribution, 167
  of death, 167
airways
  surface area, 23
  volume, 22
Airy’s equation, 241
alcohol
  in the blood, 185
algorithm, 29
allele, 146, 165
alligator, 101
alternating series, 240
alveoli, 17
analytic, 214
  approach, 29
annuity, 74
anti-differentiation, 49
antiderivative, 47, 110
  table of, 49
applications
  of integration, 61
approximation
  left endpoint, 224
  linear, 36, 200
  right endpoint, 224
Archimedes, 4
area
  as a function, 39
  circle, 6
  of planar region, 27
  of simple shapes, 1
  parallelogram, 2
  polygon, 3
  rectangle, 2
  triangle, 2
  true, 35
average, 234
  mass density, 86
  of probability distribution, 137
  weighted, 234
average value
  of a function, 76, 161
bacterial
  motion, 150
balance
  energy, 188
  mass, 186
bank
  interest rate, 74
bell
  curve, 145
Bernoulli trial, 140
bifurcate, 18
bin, 166, 233
binomial
  coefficient, 142
  distribution, 140, 143
  theorem, 142
birth, 71, 178
blood alcohol, 185
branch
  daughter, 18
  parent, 18
<table>
<thead>
<tr>
<th>Term</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bronchial tubes</td>
<td>17</td>
</tr>
<tr>
<td>calculus</td>
<td></td>
</tr>
<tr>
<td>motivation for</td>
<td>xvii</td>
</tr>
<tr>
<td>carrying capacity</td>
<td>191</td>
</tr>
<tr>
<td>center</td>
<td></td>
</tr>
<tr>
<td>of mass</td>
<td>81, 85, 133, 137, 157, 228</td>
</tr>
<tr>
<td>centroid</td>
<td>122</td>
</tr>
<tr>
<td>chain rule</td>
<td>107</td>
</tr>
<tr>
<td>change</td>
<td></td>
</tr>
<tr>
<td>net</td>
<td>66, 71</td>
</tr>
<tr>
<td>total</td>
<td>66</td>
</tr>
<tr>
<td>chemical kinetics</td>
<td>185</td>
</tr>
<tr>
<td>chromosomes</td>
<td>146</td>
</tr>
<tr>
<td>circadian rhythm</td>
<td>72</td>
</tr>
<tr>
<td>cohort</td>
<td>196, 197</td>
</tr>
<tr>
<td>coin</td>
<td></td>
</tr>
<tr>
<td>fair</td>
<td>134</td>
</tr>
<tr>
<td>toss</td>
<td>136, 165</td>
</tr>
<tr>
<td>combination</td>
<td>237</td>
</tr>
<tr>
<td>comparison</td>
<td></td>
</tr>
<tr>
<td>integral and series</td>
<td>206</td>
</tr>
<tr>
<td>integrals</td>
<td>205</td>
</tr>
<tr>
<td>tests</td>
<td>239</td>
</tr>
<tr>
<td>completing the square</td>
<td>117</td>
</tr>
<tr>
<td>conservation</td>
<td></td>
</tr>
<tr>
<td>of energy</td>
<td>188</td>
</tr>
<tr>
<td>of mass</td>
<td>186</td>
</tr>
<tr>
<td>converge</td>
<td>14</td>
</tr>
<tr>
<td>convergence</td>
<td>199</td>
</tr>
<tr>
<td>of series</td>
<td>200</td>
</tr>
<tr>
<td>tests for</td>
<td>201, 206, 237</td>
</tr>
<tr>
<td>convergent</td>
<td>15</td>
</tr>
<tr>
<td>coordinate</td>
<td></td>
</tr>
<tr>
<td>system</td>
<td>28</td>
</tr>
<tr>
<td>critical point</td>
<td>53</td>
</tr>
<tr>
<td>cumulative</td>
<td></td>
</tr>
<tr>
<td>function</td>
<td>84, 136, 154, 155, 235</td>
</tr>
<tr>
<td>data</td>
<td>133</td>
</tr>
<tr>
<td>set</td>
<td>133</td>
</tr>
<tr>
<td>decay</td>
<td></td>
</tr>
<tr>
<td>radioactive</td>
<td>162</td>
</tr>
<tr>
<td>rate</td>
<td>185</td>
</tr>
<tr>
<td>definite</td>
<td></td>
</tr>
<tr>
<td>integral</td>
<td>37, 43</td>
</tr>
<tr>
<td>density</td>
<td>61, 82</td>
</tr>
<tr>
<td>probability</td>
<td>153</td>
</tr>
<tr>
<td>dice</td>
<td>139</td>
</tr>
<tr>
<td>differential</td>
<td></td>
</tr>
<tr>
<td>equation</td>
<td>177</td>
</tr>
<tr>
<td>notation</td>
<td>107, 108</td>
</tr>
<tr>
<td>differential equation</td>
<td></td>
</tr>
<tr>
<td>linear</td>
<td>184</td>
</tr>
<tr>
<td>nonlinear</td>
<td>192</td>
</tr>
<tr>
<td>displacement</td>
<td>62</td>
</tr>
<tr>
<td>distribution</td>
<td></td>
</tr>
<tr>
<td>binomial</td>
<td>140</td>
</tr>
<tr>
<td>frequency</td>
<td>233</td>
</tr>
<tr>
<td>Gaussian</td>
<td>145</td>
</tr>
<tr>
<td>grade</td>
<td>133, 137</td>
</tr>
<tr>
<td>normal</td>
<td>145</td>
</tr>
<tr>
<td>uniform</td>
<td>169, 174</td>
</tr>
<tr>
<td>diverge</td>
<td>14, 201</td>
</tr>
<tr>
<td>divergence</td>
<td>199</td>
</tr>
<tr>
<td>of series</td>
<td>200</td>
</tr>
<tr>
<td>divergent</td>
<td>15</td>
</tr>
<tr>
<td>dummy</td>
<td></td>
</tr>
<tr>
<td>variable</td>
<td>40</td>
</tr>
<tr>
<td>emptying</td>
<td></td>
</tr>
<tr>
<td>container</td>
<td>186</td>
</tr>
<tr>
<td>time</td>
<td>191</td>
</tr>
<tr>
<td>endpoints</td>
<td>30, 113</td>
</tr>
<tr>
<td>energy</td>
<td></td>
</tr>
<tr>
<td>balance</td>
<td>188</td>
</tr>
<tr>
<td>conservation</td>
<td>188</td>
</tr>
<tr>
<td>kinetic</td>
<td>188</td>
</tr>
<tr>
<td>potential</td>
<td>188</td>
</tr>
<tr>
<td>error</td>
<td></td>
</tr>
<tr>
<td>approximation</td>
<td>216</td>
</tr>
<tr>
<td>Euler’s method</td>
<td>184</td>
</tr>
<tr>
<td>evaluate</td>
<td></td>
</tr>
<tr>
<td>a function</td>
<td>208</td>
</tr>
<tr>
<td>even</td>
<td></td>
</tr>
<tr>
<td>function</td>
<td>51</td>
</tr>
<tr>
<td>expected value</td>
<td>137</td>
</tr>
<tr>
<td>experiment</td>
<td>134</td>
</tr>
<tr>
<td>coin-toss</td>
<td>137</td>
</tr>
<tr>
<td>repeated</td>
<td>134</td>
</tr>
<tr>
<td>exponential</td>
<td>35</td>
</tr>
</tbody>
</table>
decaying, 202
function, 214
growth, 19, 180
eye
color, 146, 147
factorial, 238
notation, 236
factoring
denominator, 117
failure, 140
fair
dice, 139
falling object, 181
first-order
differential equation, 178
force
frictional, 181
of gravity, 181
formulae
areas, 25
volumes, 25
fractals, 18
frequency, 73, 136
friction, 181
frictional
coefficient, 181
fulcrum, 226
function
bounded, 37
continuous, 37
even, 51
inverse, 53
Fundamental Theorem of Calculus, 40, 41, 43, 47, 62, 155, 216
Gauss, 11
formula, 11, 12
Gaussian
distribution, 145
gene, 146
genetics, 146
genotype, 146
geometric
series, 10, 209, 240
series, finite, 13
series, finite, 200
series, infinite, 201
Gompertz, 196
grade
distribution, 137, 232
growth
density dependent, 191
exponential, 19, 75, 180
logistic, 191
population, 197
self-similar, 18
unlimited, 179, 191
growth rate
intrinsic, 191
per capita, 179
Hanoi
tower of, 8
Hardy-Weinberg, 146
harmonic
series, 201, 206, 211, 237
height
distribution, 166
higher order terms, 200
hormone
level of, 72
hypotenuse, 121
implication, 237
improper
integral, 58, 162–164, 203
income
stream, 74
induction, 221
mathematical, 12
infinite
series, 14, 200
initial
condition, 179
initial value, 178
problem, 178, 179
integral, 110
applications of, 61
converges, 202, 204
definite, 31, 33, 37, 40, 43, 110
definite, properties of, 44
diverges, 202, 203
does not exist, 57, 121
exists, 163
improper, 58, 76, 162–164, 199, 202, 206
indefinite, 110, 192
integrand, 110
integration, 33
by partial fractions, 124
by parts, 107, 124, 126
by substitution, 111
constant, 111
numerical, 162
interest
 compounded, 75
 rate, 74
inverse function, 53
inverse trigonometric functions, 121

keratocyte, 84
kinetic
energy, 188
Kulesa
 Paul, 101

leaf
area of, 33
leaking
container, 186
Leibniz, 240
length
of curve, 81, 96
of straight line, 96
limit, 29
linear approximation, 36
logistic
equation, 191
growth, 191
lung
branching, 16
human, 22
Maple, 107
mass
balance, 186
conservation, 186
density, 82, 165
discrete, 165
mass distribution
continuous, 82
discrete, 82
Mathematica, 107
mating
table, 148
maximum, 53, 55
mean, 76, 133, 137, 153, 158, 161, 234
continuous probability, 154
decay time, 162, 164
of a distribution, 157
of a probability distribution, 106
of binomial distribution, 144
measurement, 133
median, 87, 133, 158, 161, 236
continuous probability, 154
decay time, 162, 164
micron(\(\mu m\)), 84
minimum, 53, 55
model
derivation of, 186
modeling, 177
Mogilner
 Alex, 84
moment, 171
j’th, 139
first, 172
of a distribution, 171
of distribution, 139
of mass, 227
second, 139, 172
zero’th, 172
mortality, 178
age distribution, 167
constant, 178
Gompertz law of, 196
nonconstant, 196
motion
uniform, 63
uniformly accelerated, 63
multiplication
principle, 140
Murray, James D., 101
net change, 67
Newton’s law, 177
nonlinear
  differential equation, 189
normalization, 145, 155
  constant, 155, 163
numerical
  approach, 29
  method, 184
observation, 133
ODE, 177
oscillation, 73
outcome
  of experiment, 134
partial fractions, 118, 192, 231
partial sums, 15, 200
PDE, 177
pendulum, 177
perfect square, 117
period, 73
permutation, 142, 237
phenotype, 146
pi(\pi)
  approximation for, 212
  definition of, 5
polygon, 3
polynomials, 208
population
  growth, 178, 197
  sustainable, 195
potential
  energy, 188
power
  series, 199
Preface, xvii
present value, 75, 76
probability
  applications of, 161
  continuous, 153, 165
  cumulative, 136
  density, 154
  discrete, 165
  discrete, rules of, 135
  empirical, 134, 136
  symmetric, 160
  theoretical, 135
product rule
  for derivatives, 126
production, 71
progression
  geometric, 20
  mathematical, 19
Pythagorean
  theorem, 96
  triangle, 121
radioactive
  decay, 162
radioactive decay
  cumulative, 164
raindrops, 169
random
  event, 134
  variable, 135
  walk, 150
random variable
  continuous, 153
  discrete, 153
rate
  birth, 178
  mortality, 178
  of change, 67
  production, 72
  removal, 72
ratio
  test, 238
rational
  function, 124, 231
rectangle
  height of, 30
rectangular strips, 28, 43
recursion relation, 19
removal, 71
replicate, 136
rescale, 145
Riemann
  sum, 28–31, 33, 40
rule
  chain, 116
rules
<table>
<thead>
<tr>
<th>Term</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>iterated</td>
<td>23</td>
</tr>
<tr>
<td>savings account</td>
<td>74</td>
</tr>
<tr>
<td>scaled</td>
<td></td>
</tr>
<tr>
<td>equation</td>
<td>192</td>
</tr>
<tr>
<td>scaling</td>
<td>192</td>
</tr>
<tr>
<td>secant</td>
<td>230</td>
</tr>
<tr>
<td>separation</td>
<td></td>
</tr>
<tr>
<td>of variables</td>
<td>65, 177–179, 182, 189, 197</td>
</tr>
<tr>
<td>series</td>
<td></td>
</tr>
<tr>
<td>(p)</td>
<td>208</td>
</tr>
<tr>
<td>alternating</td>
<td>240</td>
</tr>
<tr>
<td>comparison tests</td>
<td>239</td>
</tr>
<tr>
<td>converges</td>
<td>200</td>
</tr>
<tr>
<td>divergent</td>
<td>238</td>
</tr>
<tr>
<td>diverges</td>
<td>200, 208</td>
</tr>
<tr>
<td>finite geometric</td>
<td>13</td>
</tr>
<tr>
<td>geometric</td>
<td>10, 13, 200, 209, 239</td>
</tr>
<tr>
<td>harmonic</td>
<td>201, 206, 211, 237, 238</td>
</tr>
<tr>
<td>infinite</td>
<td>14, 199, 200</td>
</tr>
<tr>
<td>operations on</td>
<td>240</td>
</tr>
<tr>
<td>Taylor</td>
<td>199, 209</td>
</tr>
<tr>
<td>term by term integration</td>
<td>210</td>
</tr>
<tr>
<td>Sigma</td>
<td></td>
</tr>
<tr>
<td>notation</td>
<td>9</td>
</tr>
<tr>
<td>size distribution</td>
<td>169</td>
</tr>
<tr>
<td>sketching</td>
<td></td>
</tr>
<tr>
<td>antiderivative</td>
<td>53</td>
</tr>
<tr>
<td>solids</td>
<td></td>
</tr>
<tr>
<td>of revolution</td>
<td>90, 91</td>
</tr>
<tr>
<td>solution</td>
<td></td>
</tr>
<tr>
<td>curves</td>
<td>190</td>
</tr>
<tr>
<td>of initial value problem</td>
<td>179</td>
</tr>
<tr>
<td>qualitative</td>
<td>68</td>
</tr>
<tr>
<td>quantitative</td>
<td>68</td>
</tr>
<tr>
<td>to ODE</td>
<td>180</td>
</tr>
<tr>
<td>spreadsheet</td>
<td>23, 29, 190</td>
</tr>
<tr>
<td>standard deviation</td>
<td>138, 172, 173</td>
</tr>
<tr>
<td>steady state</td>
<td>66, 184, 195</td>
</tr>
<tr>
<td>step</td>
<td></td>
</tr>
<tr>
<td>function</td>
<td>138, 235</td>
</tr>
<tr>
<td>strips</td>
<td></td>
</tr>
<tr>
<td>area of</td>
<td>28</td>
</tr>
<tr>
<td>rectangular</td>
<td>28, 43</td>
</tr>
<tr>
<td>substitution</td>
<td>107, 111</td>
</tr>
<tr>
<td>examples</td>
<td>113</td>
</tr>
<tr>
<td>trigonometric</td>
<td>118, 123</td>
</tr>
<tr>
<td>success</td>
<td>140</td>
</tr>
<tr>
<td>sum</td>
<td></td>
</tr>
<tr>
<td>geometric</td>
<td>35</td>
</tr>
<tr>
<td>of (N) cubes</td>
<td>12</td>
</tr>
<tr>
<td>of (N) integers</td>
<td>11</td>
</tr>
<tr>
<td>of (N) squares</td>
<td>12</td>
</tr>
<tr>
<td>of square integers</td>
<td>32</td>
</tr>
<tr>
<td>Riemann</td>
<td>29, 30, 40</td>
</tr>
<tr>
<td>summation</td>
<td></td>
</tr>
<tr>
<td>index</td>
<td>9</td>
</tr>
<tr>
<td>notation</td>
<td>9</td>
</tr>
<tr>
<td>sums</td>
<td></td>
</tr>
<tr>
<td>partial</td>
<td>15, 200</td>
</tr>
<tr>
<td>surface area</td>
<td></td>
</tr>
<tr>
<td>cylinder</td>
<td>6</td>
</tr>
<tr>
<td>survival</td>
<td></td>
</tr>
<tr>
<td>probability</td>
<td>168</td>
</tr>
<tr>
<td>tangent line</td>
<td>200</td>
</tr>
<tr>
<td>Taylor polynomial</td>
<td>209</td>
</tr>
<tr>
<td>Taylor series</td>
<td>199, 209</td>
</tr>
<tr>
<td>for (\cos(x))</td>
<td>216</td>
</tr>
<tr>
<td>for (\sin(x))</td>
<td>214</td>
</tr>
<tr>
<td>for (e^x)</td>
<td>213</td>
</tr>
<tr>
<td>teeth</td>
<td>99</td>
</tr>
<tr>
<td>temperature</td>
<td>67</td>
</tr>
<tr>
<td>terminal velocity</td>
<td>180</td>
</tr>
<tr>
<td>torque</td>
<td>226, 227</td>
</tr>
<tr>
<td>tree</td>
<td></td>
</tr>
<tr>
<td>growth</td>
<td>68</td>
</tr>
<tr>
<td>structure</td>
<td>18</td>
</tr>
<tr>
<td>trial</td>
<td></td>
</tr>
<tr>
<td>Bernoulli</td>
<td>140</td>
</tr>
<tr>
<td>triangle</td>
<td></td>
</tr>
<tr>
<td>Pythagorean</td>
<td>121</td>
</tr>
<tr>
<td>trigonometric</td>
<td>120</td>
</tr>
<tr>
<td>trifurcate</td>
<td>18</td>
</tr>
<tr>
<td>trigonometric</td>
<td></td>
</tr>
<tr>
<td>identities</td>
<td>118</td>
</tr>
<tr>
<td>substitution</td>
<td>118</td>
</tr>
<tr>
<td>unbiased, 134, 136</td>
<td></td>
</tr>
<tr>
<td>unbounded, 134</td>
<td></td>
</tr>
<tr>
<td>function, 57</td>
<td></td>
</tr>
</tbody>
</table>
undefined
  function, 58
units, 7

variance, 138, 153, 172
  continuous probability, 154
velocity, 62
  terminal, 66, 180, 184
volume
  cube, 6
  cylinder, 7, 90
  cylindrical shell, 7
disk, 90
  of solids, 81
  rectangular box, 6
  shell, 90
  simple shapes, 6
  sphere, 7
  spherical shell, 7

zygote, 147