# Math 141 Course Notes
Austin Christian

## Contents

1 **Functions of Several Variables** 3
   1.1 Definitions and Vocabulary .............................. 3
   1.2 Partial Derivatives .................................... 7
   1.3 Optimization ........................................ 11
   1.4 Lagrange Multipliers .................................. 17
   1.5 The Method of Least Squares .......................... 22

2 **Matrices** 26
   2.1 Systems of Linear Equations with Unique Solutions .......... 26
   2.2 General Systems of Linear Equations ........................ 32
   2.3 Arithmetic Operations on Matrices .......................... 38
   2.4 The Inverse of a Matrix ................................ 43
   2.5 The Gauss-Jordan Method for Matrix Inversion ................ 47
   2.6 Input-Output Analysis .................................. 49

3 **Sets and Counting** 53
   3.1 Sets .............................................. 53
   3.2 A Fundamental Principle of Counting ....................... 57
   3.3 Venn Diagrams and Counting .............................. 60
   3.4 The Multiplication Principle ............................. 62
   3.5 Permutations and Combinations ............................ 64
   3.6 Further Counting Techniques .............................. 68
      3.6.1 Stars and Bars .................................. 68
      3.6.2 The MISSISSIPPI Problem .......................... 70
      3.6.3 Increasing Order Problems ........................ 72
      3.6.4 Some Additional Examples ......................... 73

4 **Probability** 75
   4.1 Experiments, Outcomes, Sample Spaces, and Events .......... 75
   4.2 Assignment of Probabilities ................................ 78
   4.3 Calculating Probabilities of Events ........................ 83
      4.3.1 Experiments with Equally Likely Outcomes ............ 83
      4.3.2 The Complement Rule ................................ 84
   4.4 Conditional Probability and Independence ................... 86
   4.5 Tree Diagrams ....................................... 92
   4.6 Bayes’ Theorem, Natural Frequencies ...................... 95
5 Probability and Statistics

5.1 Visual Representations of Data ........................................ 98
5.2 Frequency and Probability Distributions .......................... 101
5.3 Binomial Trials ........................................................... 107
5.4 The Mean .............................................................. 110
5.5 The Variance and Standard Deviation ............................. 116
1 Functions of Several Variables

In this chapter we draw on discussions and examples found in [1] and [2].

1.1 Definitions and Vocabulary

Much of the vocabulary and lingo surrounding functions of several variables is a simple extension of the vocabulary we have for single-variable functions, but we will also add a few new terms.

Recall that a function \( f \) of a single variable is a rule that assigns a unique real number \( f(x) \) to each number \( x \) in some set \( D \) in \( \mathbb{R} \). (As you’ve probably seen, we use \( \mathbb{R} \) to mean all real numbers. We’ll use \( \mathbb{R}^2 \) to indicate the \( xy \)-plane.) We call the set \( D \) where \( f \) is defined the domain of \( f \), and we call the values \( f(x) \) that come out of our rule the range of \( f \). For example, \( f(x) = 1/x \) defines a function on the domain \( D \), where \( D \) is all nonzero real numbers. The story is similar when we add extra variables.

**Definition.** A function \( f \) of two variables is a rule that assigns a unique real number \( f(x,y) \) to each point \((x,y)\) in some set \( D \) in the \( xy \)-plane \( \mathbb{R}^2 \). A function \( f \) of \( n \) variables is a rule that assigns a unique real number \( f(x_1, \ldots, x_n) \) to each point \((x_1, \ldots, x_n)\) in some set \( D \) in \( \mathbb{R}^n \).

If the second part of the above definition scares you, don’t worry too much; we’ll almost always restrict our attention to functions of two variables.

**Example 1.1.** Colleen is planning a wedding ceremony, with a reception to follow. The venue charges a base rental fee of $1,000, charges $1.50 for each chair they set up at the ceremony, and charges $2.00 for each table setting at the reception.

(a) Write down a function \( C(x, y) \) which gives the venue cost for \( x \) guests at the ceremony and \( y \) guests at the reception.

(b) What is the domain of the cost function?

(c) Determine the venue cost if 200 guests attend the ceremony and 180 guests stay for the reception.

(d) What is the difference in cost if one extra guest attends the ceremony and reception?

(Solution.)

---

We’re restricting our attention to real-valued functions of a real variable. In general, a function can accept inputs and produce outputs that aren’t real numbers.
(a) Since each guest at the ceremony costs $1.50, the cost of \(x\) guests at the ceremony will be $1.50x. Similarly, the cost of \(y\) guests at the reception will be $2y. Adding these costs to the base rental fee gives

\[
C(x, y) = 1,000 + 1.50x + 2y.
\]

Notice that the total venue cost is a function of two variables: the number of guests at the ceremony and the number of guests at the reception.

(b) Certainly neither \(x\) nor \(y\) may be negative. It also seems safe to assume that our guest total will be an integer at both the ceremony and the reception — that is, we won’t have, say, 121.34 guests at the ceremony. So our domain is all of those points \((x, y)\) where \(x\) and \(y\) are nonnegative integers.

(c) With 200 guests at the ceremony and 180 guests at the reception, the total venue cost will be

\[
C(200, 180) = 1,000 + 1.50 \cdot 200 + 2 \cdot 180 = 1,000 + 300 + 360 = 1,660.
\]

(d) An extra guest at the ceremony and reception will mean an extra chair ($1.50) and an extra table setting ($2), for a total extra cost of $3.50. We could make this a lot harder than it has to be by noticing that an extra guest at the ceremony means we replace \(x\) with \(x + 1\) and an extra guest at the reception means we replace \(y\) with \(y + 1\). The difference in cost is then

\[
C(x + 1, y + 1) - C(x, y) = (1,000 + 1.50(x + 1) + 2(y + 1)) - (1,000 + 1.50x + 2y) = 3.50.
\]

Unfortunately some of the applications we deal with are not so straightforward as this one, and this messier algebraic approach will become necessary. As a final note we point out that $3.50 is just the monetary cost of an extra guest. Showing up uninvited will also land one on the bride’s bad side, which is a bad place to be indeed.

□

After you first learned the definition of a single-variable function, you probably spent a great deal of time studying the plots or graphs of these functions. We obtain these plots as follows: if \(f\) is a function of a single variable, we plot all the points \((x, y)\) in the plane which satisfy the relationship \(y = f(x)\). For example, our function above was \(f(x) = 1/x\), defined on the domain \(D\) consisting of all nonzero real numbers. The plot of this function is given in Figure 1a.

Little changes when we add an extra variable. We still want to plot our functions, only now we need an extra dimension. A single-variable function required two dimensions — that
is, the plane — to plot; plots of two-variable functions live in three-dimensional space. Now
instead of plotting solutions to the equation \( y = f(x) \) we’ll plot solutions to the equation
\( z = f(x, y) \), where \((x, y, z)\) is an equation in 3-space. For example, Figure 1b gives the plot of
the cost function \( C(x, y) \) found in our above example. Note that the height of the resulting
surface increases as we increase either \( x \) or \( y \), and it increases more quickly as we increase \( y \)
than it does as we increase \( x \). (Why is this the case?)

With the extra dimension of freedom we see a greater variety of behaviors from plots of
two variables, but the plots generally behave in a manner that you would expect from your
experience with single-variable functions. For example, the plot of the function \( f(x, y) = x^2 + y^2 \)
obtained by adding two quadratic terms gives a surface we call a **paraboloid**, which
has the shape of an upward-opening parabola in both the \( x \)- and \( y \)-direction, as seen in
Figure 2a. On the other hand, the function \( f(x) = x^2 - y^2 \) is obtained by subtracting one
quadratic term from another. This gives a downward-opening parabola in the \( y \)-direction
and an upward opening parabola in the \( x \)-direction to form the **saddle** seen in Figure 2b.

We’ll end this section with a consideration we didn’t have for functions of a single variable:
level curves. Remember that we think of functions as producing some output depending on the input we’ve given. We’re often interested in knowing which input values will produce a particular output value. That is, if \( z_0 \) is a fixed real number, we want to know which points \((x, y)\) satisfy \( z_0 = f(x, y) \). For instance, in our above example about wedding guests, perhaps Colleen would like to know what combinations of ceremony guests and reception guests she can afford with a venue budget of $1,200. This type of problem comes up a lot, and we have a name for the plots that are produced.

**Definition.** Given a function \( f \) of two variables and a fixed value \( z_0 \), we call the intersection of the surface \( z = f(x, y) \) and the plane \( z = z_0 \) the **level curve of height** \( z_0 \) for the function \( f \).

As an example, Figure 3a shows the saddle \( z = x^2 - y^2 \) that we plotted earlier along with several level curves. These level curves are projected onto the \( xy \)-plane in Figure 3b; we call a plot of the level curves a **contour plot** of \( f \).

![Level curves and contour plot](image)

(a) The saddle \( z = x^2 - y^2 \) with some level curves. (b) A contour plot of the saddle \( z = x^2 - y^2 \).

**Figure 3**

**Example 1.2.** If \( f \) is a function that tells us the cost associated with some variables \( x \) and \( y \) (that is, \( f \) is a cost function), we sometimes call the level curves of \( f \) **isocost curves** or **isocost lines**. Referring back to Example 1.1, plot the isocost lines for Colleen’s wedding associated to the budgets $1,100, $1,200, and $1,300.

*(Solution.)* In Example 1 we found that the cost function for Colleen’s wedding is

\[
C(x, y) = 1,000 + 1.50x + 2y,
\]

where \( x \) represents the number of guests at the ceremony and \( y \) represents the number of guests at the reception. The isocost lines are then given by

\[
C_0 = 1,000 + 1.50x + 2y,
\]

where \( C_0 \) is the cost associated with a budget of \( \$C_0 \).
where $C_0$ is a fixed cost. For a budget of $1,100 this gives

$$1,100 = 1,000 + 1.50x + 2y \quad \Rightarrow \quad 100 = 1.50x + 2y.$$ 

Similarly, the budgets of $1,200 and $1,300 give linear equations

$$200 = 1.50x + 2y \quad \text{and} \quad 300 = 1.50x + 2y,$$

respectively. Figure 4 gives a plot of these three isocost lines: In this plot, the blue line is the isocost line for a budget of $1,100 while the orange and green lines correspond to budgets of $1,200 and $1,300, respectively. The interpretation for this plot is that these are the lines along which the budget is constant. For example, with a budget of $1,300 Colleen could invite 200 guests to the ceremony and no one to the reception, 150 guests to the reception and no one to the ceremony, or (preferably) some combination that is found along the green line.

1.2 Partial Derivatives

Now that we’ve become comfortable with the basic vocabulary surrounding functions of multiple variables, we would like to extend our knowledge of calculus to these types of functions. As with graphs, the extra variable(s) we now have will lead to a richer discussion.

Recall that the derivative of a single-variable function $f$ is intended to measure the extent to which the output $f(x)$ is affected by small changes in the input $x$. We make this measurement by computing the slope of the tangent line to the graph of $f$ at the point $(x, f(x))$ and find that the derivative is given by

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h},$$

when the limit on the right exists. When we move to a function of two variables we have two inputs which might undergo small changes, and we would like to measure the effects of these
changes on the output $f(x, y)$. The geometric interpretation is that graphs of functions of two variables don’t have tangent lines, but instead have tangent planes, as you see in Figure 5. Instead of having just one interesting derivative, functions of two variables have several, the most important of which are the **partial derivatives** with respect to $x$ and $y$.

The partial derivative of $f$ with respect to $x$ is intended to measure the effect of a small change in $x$ on the output $f(x, y)$, so it makes sense to mimic the definition seen in (1). Indeed, this is the definition of the partial derivative:

$$\frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h},$$

when the limit on the right exists. (The funny $\partial$ symbol often serves as a helpful reminder that our function depends on more than one variable. If we used $d$ we might forget that there’s also a $y$-derivative.) We can write down a similar definition for the partial derivative of $f$ with respect to $y$, but it turns out that partial derivatives are much easier to compute than these definitions let on. Indeed, we’ll give a less formal definition that makes these computations clear.

**Definition.** Let $f$ be a function of two variables, and let $(x_0, y_0)$ be a fixed point in the domain of $f$. Then the **partial derivative of $f$ with respect to $x$** at $(x_0, y_0)$ is given by treating $y$ as a constant:

$$\frac{\partial f}{\partial x}(x_0, y_0) = \frac{d}{dx}(f(x, y_0))|_{x_0}.$$  

Similarly, the **partial derivative of $f$ with respect to $y$** at $(x_0, y_0)$ is given by treating $x$ as a constant:

$$\frac{\partial f}{\partial y}(x_0, y_0) = \frac{d}{dy}(f(x_0, y))|_{y_0}.$$  

For brevity, we sometimes also use the notation $f_x$ and $f_y$ for the partial derivatives of $f$. If $f$ is a function of more than two variables, the definitions are analogous.

This definition should become even clearer by way of examples.
Example 1.3. Find $f_x$ and $f_y$, where $f$ is given by:

(a) $f(x, y) = x^2 + 3y$;   
(b) $f(x, y) = x/y$;   
(c) $f(x, y) = (x + xy + 2y)^3$.

(Solution.)

(a) When computing the $x$-partial derivative of $f$, we treat $y$ as a constant:

$$\frac{\partial f}{\partial x} = \frac{d}{dx}(x^2 + 3y) = 2x + 0 = 2x.$$

Similarly, $f_y = \frac{d}{dy}(x^2 + 3y) = 3$.

(b) The $x$-partial is easy here; since we treat $y$ as a constant, $f$ looks like a constant multiple of $x$, so we have $f_x = 1/y$. For the $y$-partial we treat $x$ as a constant, and it might be helpful to remember that we can “pull constants out” when computing derivatives:

$$f_y = \frac{d}{dy} \left( \frac{x}{y} \right) = x \cdot \frac{d}{dy} \left( \frac{1}{y} \right) = x \cdot \frac{-1}{y^2} = \frac{-x}{y^2}.$$

(c) This last example will require the chain rule. We have

$$f_x = \frac{d}{dx} \left( (x + xy + 2y)^3 \right) = 3(x + xy + 2y)^2 \cdot \frac{d}{dx} (x + xy + 2y)$$

$$= 3(x + xy + 2y)^2 (1 + y).$$

Similarly,

$$f_y = \frac{d}{dy} \left( (x + xy + 2y)^3 \right) = 3(x + xy + 2y)^2 \cdot \frac{d}{dy} (x + xy + 2y)$$

$$= 3(x + xy + 2y)^2 (x + 2).$$

So even when these computations get messy they are usually just a steady application of our tools for single-variable derivatives.

□

After learning all the rules that go along with differentiation, it’s easy to lose sight of our original intentions: to measure the way that the output $f(x)$ depends on the input $x$. Suppose $a$ is some very small (but nonzero) number, and we want to see how different $f(x + a)$ is from $f(x)$. That is, we want to know $f(x + a) - f(x)$. Now since

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h},$$
we should have
\[ f'(x) \approx \frac{f(x + a) - f(x)}{a}, \]
if \(a \neq 0\) is small enough. Multiplying both sides by \(a\), we see that
\[ f(x + a) - f(x) \approx af'(x). \]

We can play a similar game with partial derivatives:

If \(a\) and \(b\) are small numbers, then \(f\) has local linear approximations given by
\[ f(x + a, y) - f(x, y) \approx a \cdot \frac{\partial f}{\partial x}(x, y) \]
and
\[ f(x, y + b) - f(x, y) \approx b \cdot \frac{\partial f}{\partial y}(x, y). \]

If these small changes are made simultaneously, we have the linear approximation
\[ f(x + a, y + b) - f(x, y) \approx a \cdot \frac{\partial f}{\partial x}(x, y) + b \cdot \frac{\partial f}{\partial y}(x, y). \] (2)

Notice that this approximation includes the previous two by allowing \(a = 0\) or \(b = 0\).

**Example 1.4.** Consider the production function \(f(x, y) = 60x^{3/4}y^{1/4}\), which gives the number of units of goods produced when \(x\) units of labor and \(y\) units of capital are used. Find the production level when 81 units of labor and 16 units of capital are used, and use this to approximate the production level when 80 units of labor and 18 units of capital are used.

*(Solution.)* We start by computing \(f(81, 16)\). We have
\[ f(81, 16) = 60(81)^{3/4}(16)^{1/4} = 60(3)^{3}(2)^{1} = 60 \cdot 27 \cdot 2 = 3,240 \text{ units}. \]

To approximate \(f(80, 18)\) using (2), we’ll need to compute \(f_x(81, 16)\) and \(f_y(81, 16)\). (These quantities are known as the marginal productivity of labor and the marginal productivity of capital, respectively.) We have
\[ \frac{\partial f}{\partial x} = 60 \cdot \frac{3}{4}x^{-1/4}y^{1/4}, \]
so
\[ \frac{\partial f}{\partial x}(81, 16) = 60 \cdot \frac{3}{4}(81)^{-1/4}(16)^{1/4} = 60 \cdot \frac{3}{4} \cdot \frac{1}{3} \cdot 2 = 30. \]

Similarly,
\[ \frac{\partial f}{\partial y} = 60 \cdot \frac{1}{4}x^{3/4}y^{-3/4}, \]
so
\[
\frac{\partial f}{\partial y}(81, 16) = 60 \cdot \frac{1}{4}(81)^{3/4}(16)^{-3/4} = 60 \cdot \frac{1}{4}(3)^3(2)^{-3} = 15 \cdot \frac{27}{8} = 50.625.
\]

We may apply (2) to see that
\[
f(80, 18) - f(81, 16) \approx (-1)(30) + (2)(50.625) = -30 + 101.25 = 71.25.
\]

So this change to our labor-capital arrangement will see an increase in production of about 71 units of goods. In particular, the production level will rise to about 3,311.25 units. □

Finally we note that just as with usual derivatives, we can take higher-order partial derivatives, or derivatives of derivatives. Some notation for these higher-order derivatives includes
\[
\frac{\partial^3 f}{\partial x^3} = \frac{\partial}{\partial x} \left( \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial x} \right) \right), \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) = f_{yx}, \quad \text{and} \quad f_{xx} = \frac{\partial^2}{\partial x^2}.
\]

We call the term in the middle a mixed partial derivative because it involves taking a partial derivative with respect to each variable. These can look a little concerning at first glance, because who wants to have to remember whether \(f_{xy}\) means to take the \(x\)-partial and then the \(y\)- or the other way around? Thankfully, the following theorem saves us from this concern in almost all cases we’ll consider.

**Theorem 1** (Clairaut’s Theorem.) If \(f\) is a function of two variables and \(f_{xy}\) and \(f_{yx}\) are both continuous, then \(f_{xy} = f_{yx}\).

**Example 1.5.** Verify Theorem 1 for the function \(f(x, y) = (2x + 3y)^2\).

**Solution.** We have
\[
f_{xy}(x, y) = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right) = \frac{\partial}{\partial y} \left( 2(2x + 3y)(2) \right) = \frac{\partial}{\partial y} \left( 4(2x + 3y) \right) = 12.
\]

At the same time,
\[
f_{yx}(x, y) = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) = \frac{\partial}{\partial x} \left( 2(2x + 3y)(3) \right) = \frac{\partial}{\partial x} \left( 6(2x + 3y) \right) = 12.
\]

So indeed \(f_{xy} = f_{yx}\). □

### 1.3 Optimization

One of the most important (and applicable) uses of the calculus of single-variable functions is to find (relative) maxima or minima of a function. For example, we can use calculus to find the relative maxima of a revenue function \(R(x)\) or the relative minima of a cost function...
We usually do this via a two-step process. First we use the first derivative test to identify candidates for relative extrema; we find those points $x$ where $f'(x) = 0$. We then use the second derivative test to check the concavity of our function near these candidates. If $f'(x) = 0$ and $f''(x) > 0$, then $f$ is concave up (or convex) at $x$, and $f(x)$ is a relative minimum for $f$. If on the other hand $f''(x) < 0$, then $f$ is concave down (or concave) at $x$, and $f(x)$ is a relative maximum for $f$. The second derivative test is inconclusive when $f''(x) = 0$, as is the case when $f(x) = x^3$ and $x = 0$ (see Figure 6). Though two-variable functions lead to more exotic graphs than single-variable functions, we will optimize them following the same general strategy.

Though it is intuitively obvious, we start by defining what we mean by a (relative) maximum or minimum of a function of two variables:

**Definition.** Let $f$ be a function of two variables and let $(x_0, y_0)$ be a point in the domain of $f$. We say that $f(x_0, y_0)$ is a relative maximum for $f$ if there is a small circle around $(x_0, y_0)$ so that for every point $(x, y)$ in the circle and in the domain of $f$, $f(x, y) \leq f(x_0, y_0)$. Similarly, $f(x_0, y_0)$ is a relative minimum if $f(x, y) \geq f(x_0, y_0)$ for every point $(x, y)$ in the circle and in the domain of $f$. We call $f(x_0, y_0)$ a relative extremum if $f(x_0, y_0)$ is either a relative maximum or a relative minimum for $f$.

This definition may seem a little technically daunting, but there is one thing we can immediately glean: if the point $(x_0, y_0)$ gives a relative maximum (or minimum) for $f$, then this point gives a relative maximum (or minimum) “in the $x$-direction” and “in the $y$-direction”.

To see what we mean, take a look at Figure 7. At the intersection of the two green curves we have a relative maximum in both the $y$-direction and the $x$-direction. At the intersection of the blue and green curves we have a relative maximum in the $y$-direction, but by changing our $x$-value we may move up along the green curve to obtain a greater value. So this point is not a relative maximum for $f$, since it is not simultaneously a relative maximum for $f$.\[\text{2}\]

\[\text{2}\]This implication cannot be reversed. It’s possible that $f$ will have a maximum in the $x$- and $y$-directions, but not in some other direction.
each variable. Similarly, at the intersection of the red and green curves we have a relative maximum in the $x$-variable, but not in the $y$-variable, so this point is not a relative maximum.

Having made this connection between two-variable optimization problems and single-variable optimization, it’s helpful to remember the first-derivative test for single-variable functions: if $f$ is a single-variable function and $f$ has a relative maximum or minimum at $x_0$, then $f'(x_0) = 0$. Since a two-variable function $f$ must have a relative extremum in each direction in order to have a relative extremum at $(x_0, y_0)$, its derivatives must vanish in each direction at this point. That is:

First derivative test. If $f$ has a relative extremum at $(x_0, y_0)$ and $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ both exist, then

$$\frac{\partial f}{\partial x}(x_0, y_0) = 0 \quad \text{and} \quad \frac{\partial f}{\partial y}(x_0, y_0) = 0.$$  

If $f_x(x_0, y_0) = 0$ and $f_y(x_0, y_0) = 0$, we call $(x_0, y_0)$ a critical point for $f$. Note that being a critical point does not guarantee that $f(x_0, y_0)$ is a relative extremum for $f$.

**Example 1.6.** The plot in Figure 7 is the graph of the function $f(x,y) = 9 - (x^2 + y^2)$. The maximum value shown occurs at $(0,0)$, where $f(0,0) = 9$. Notice that

$$\frac{\partial f}{\partial x} = -2x \quad \text{and} \quad \frac{\partial f}{\partial y} = -2y,$$

so $f_x(0,0) = 0$ and $f_y(0,0) = 0$. In fact, $f_x(0,y) = 0$ for all values of $y$ and $f_y(x,0) = 0$ for all values of $x$, as seen by the two green curves, but $(0,0)$ is the only point where both partial derivatives simultaneously vanish, and thus the only point where a relative extremum could occur.

Once we’ve identified a point $(x_0, y_0)$ where both $f_x$ and $f_y$ vanish, how are we to determine whether this point is a relative maximum, relative minimum, or neither? In the single-variable case we have the second derivative test, so perhaps we can apply this to each
of $x$ and $y$? For example, in Figure 7, $f$ has negative second derivative in both the $x$- and $y$-directions. On the other hand, the function whose graph is seen in Figure 8 has a positive second $x$-partial and a negative second $y$-partial. (We call points such as the red one seen in Figure 8 saddle points.) So perhaps we can obtain a second derivative test for two-variable functions by applying the single-variable second derivative test to each of $x$ and $y$. We do have a second derivative test for two-variable functions, but unfortunately it is not as simple as we might have hoped. Because we have to consider the concavity of $f$ in every direction, the two-variable second derivative test is significantly more complicated than the single-variable version.

**Second derivative test.** Suppose $f$ is a function of two variables and that $(x_0, y_0)$ is a critical point for $f$. Suppose $f_{xx}, f_{yy},$ and $f_{xy}$ are continuous around $(x_0, y_0)$ and let

$$D = f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0).$$

If

1. $D > 0$ and $f_{xx}(x_0, y_0) > 0$, then $f$ has a relative minimum at $(x_0, y_0)$;
2. $D > 0$ and $f_{xx}(x_0, y_0) < 0$, then $f$ has a relative maximum at $(x_0, y_0)$;
3. $D < 0$, then $f$ has a saddle point at $(x_0, y_0)$;
4. $D = 0$, then the second derivative test is inconclusive.

**Example 1.7.** Locate all relative maxima, relative minima, and saddle points of

$$f(x, y) = x^3 + y^3 - 3x - 3y,$$

if there are any.

*(Solution.*) We begin by identifying the critical points of $f$ — the points where both $f_x$ and $f_y$ vanish. We have

$$\frac{\partial f}{\partial x} = 3x^2 - 3 \quad \text{and} \quad \frac{\partial f}{\partial y} = 3y^2 - 3,$$
so setting \( f_x(x, y) = 0 \) gives \( 3x^2 - 3 = 0 \), which means \( x = \pm 1 \). Similarly, setting \( f_y(x, y) = 0 \) gives \( y = \pm 1 \). This means that \( f_x \) and \( f_y \) are both zero when \( x = \pm 1 \) and \( y = \pm 1 \). This means that we have four critical points:

\[
(-1, -1), \quad (-1, 1), \quad (1, -1), \quad (1, 1).
\]

Next we compute our second partial derivatives:

\[
\frac{\partial^2 f}{\partial x^2} = 6x, \quad \frac{\partial^2 f}{\partial y^2} = 6y, \quad \text{and} \quad \frac{\partial^2 f}{\partial x \partial y} = 0.
\]

This means that

\[
D(x, y) = (6x)(6y) - 0 = 36xy.
\]

From this we may determine whether each of our critical points is a relative maximum, relative minimum, or saddle point:

<table>
<thead>
<tr>
<th>Critical Point</th>
<th>D</th>
<th>( f_{xx} )</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>((-1, -1))</td>
<td>36 &gt; 0</td>
<td>-6 &lt; 0</td>
<td>relative maximum</td>
</tr>
<tr>
<td>((-1, 1))</td>
<td>-36 &lt; 0</td>
<td>N/A</td>
<td>saddle point</td>
</tr>
<tr>
<td>((1, -1))</td>
<td>-36 &lt; 0</td>
<td>N/A</td>
<td>saddle point</td>
</tr>
<tr>
<td>((1, 1))</td>
<td>36 &gt; 0</td>
<td>6 &gt; 0</td>
<td>relative minimum</td>
</tr>
</tbody>
</table>

The conclusions are made by applying the second derivative test. Notice that when \( D < 0 \) the second derivative test does not require us to compute \( f_{xx} \). We can verify our conclusions by consulting Figure 9.
Example 1.8. A company manufactures and sells two products, I and II that sell for $10 and $9 per unit, respectively. The cost of producing $x$ units of product I and $y$ units of product II is

$$400 + 2x + 3y + 0.01(3x^2 + xy + 3y^2).$$

Find the values of $x$ and $y$ that maximize the company’s profits.

(Solution.) Since we’re hoping to optimize the company’s profits, we should first write the profit as a function of $x$ and $y$. Since

$$\text{profit} = \text{revenue} - \text{cost},$$

we may substitute (3) to find

$$P(x, y) = (10x + 9y) - (400 + 2x + 3y + 0.01(3x^2 + xy + 3y^2)) = 8x + 6y - 400 - 0.01(3x^2 + xy + 3y^2).$$

To find the critical points of this function, we first compute the partial derivatives:

$$P_x(x, y) = 8 - 0.01(6x + y), \quad P_y(x, y) = 6 - 0.01(x + 6y).$$

For $P_x$ to vanish we must have $8 = 0.01(6x + y)$, so

$$6x + y = \frac{8}{0.01} \quad \Rightarrow \quad y = 800 - 6x.$$ 

Since we’re looking for points where $P_x$ and $P_y$ simultaneously vanish, we may substitute this into our expression for $P_y$ and solve

$$0 = 6 - 0.01(x + 6(800 - 6x))$$

for $x$. We have

$$6 = 0.01(-35x + 4800) \quad \Rightarrow \quad 600 = -35x + 4800 \quad \Rightarrow \quad -4200 = -35x \quad \Rightarrow \quad x = 120.$$ 

Because we have seen that $y = 800 - 6x$, this means that $y = 80$. So $P$ has only one critical point: $(120, 80)$. To check that this is a relative minimum for $P$, we apply the second derivative test. First, the second derivatives:

$$P_{xx}(x, y) = -0.06, \quad P_{yy}(x, y) = -0.06, \quad \text{and} \quad P_{xy} = -0.01.$$ 

So we have $D = (-0.06)(-0.06) - (-0.01)^2 = 0.0035 > 0$. Since $P_{xx} < 0$, $P$ has a relative maximum at $x = 120$ and $y = 80$. So the company should manufacture 120 units of product I and 80 units of product II, assuming all would sell. \[\square\]

Note. Recall that in single-variable calculus we found global extrema by first finding any relative extrema and then examining the end behavior of our function. A similar strategy may be employed for functions of two variables, but we won’t address this in the current course.
1.4 Lagrange Multipliers

In the previous section we focused on finding relative maxima and minima of functions, and this is certainly an applicable skill in business — we want to minimize costs while maximizing revenue. Our techniques have a troubling shortcoming, however: the maxima and minima often occur at points in our domain which are not feasible in reality. For instance, if we return to the wedding example first seen in Example 1.1 and try to minimize venue costs, we’ll conclude that Colleen’s venue cost is minimized when she invites zero guests to her wedding — an outcome that hardly seems desirable. In this section we’ll focus on what’s called constrained optimization — optimization subject to some sort of constraint.

Suppose we have a function $f(x, y)$ that we would like to maximize or minimize, but that $x$ and $y$ are bound by the requirement that $g(x, y) = 0$. For example, $x$ and $y$ might be production levels of a pair of products and we might require that $x + y - 1000 = 0$, meaning that the total production level must be 1000 units. We call the function we’re trying to optimize (in this case $f$) the objective function and we call $g(x, y) = 0$ the constraint equation.

A geometric understanding of this setup can be seen in Figure 10. The red curve represents the points $(x, y)$ that satisfy a constraint $g(x, y) = 0$ and the blue curves are level curves of a function $f(x, y)$, with the value of $f$ along each curve labeled. We would like to identify the point along the curve $g(x, y) = 0$ at which $f$ is maximized. In this instance we see that $f$ is maximized at the green dot, where we have $f(x, y) = 50$. Notice that the curve $g(x, y) = 0$ is tangent to the level curve $f(x, y) = 50$. This is not an accident or coincidence.

If $f$ has a maximum or minimum value of $c$ subject to the constraint $g(x, y) = 0$, then the curves $g(x, y) = 0$ and $f(x, y) = c$ are tangent when they meet.
Because we’ve overlooked some of the interaction between partial derivatives and the geometry of level curves, it is difficult at this point to write down any formulae that encapsulate the geometric observation we’ve made. Unfortunately, you’ll have to take the following strategy on faith\(^3\). To identify the points \((x_0, y_0)\) where \(g(x, y) = 0\) is tangent to a level curve of \(f\), we first define the following function of three variables\(^4\)

\[
F(x, y, \lambda) = f(x, y) + \lambda g(x, y). \tag{4}
\]

It is a fact that if \(g(x, y) = 0\) is tangent to a level curve of \(f\) at some point \((x_0, y_0)\), then we may find some value \(\lambda\) so that

\[
\frac{\partial F}{\partial x}(x_0, y_0, \lambda) = 0, \quad \frac{\partial F}{\partial y}(x_0, y_0, \lambda) = 0, \quad \text{and} \quad \frac{\partial F}{\partial \lambda}(x_0, y_0, \lambda) = 0.
\]

When looking for a point \((x, y)\) which maximizes or minimizes \(f\) subject to \(g(x, y) = 0\), it makes sense to look for all points where \(g(x, y) = 0\) and a level curve of \(f\) meet tangentially, and then identify which of these points optimize \(f\). We have the following strategy for optimizing \(f\) with respect to our constraint:

**The Method of Lagrange Multipliers.** Suppose we wish to optimize a function \(f(x, y)\) of two variables subject to the constraint \(g(x, y) = 0\).

1. Define a function \(F(x, y, \lambda) = f(x, y) + \lambda g(x, y)\) and set

\[
\text{(i) } \frac{\partial F}{\partial x}(x, y, \lambda) = 0, \quad \text{(ii) } \frac{\partial F}{\partial y}(x, y, \lambda) = 0, \quad \text{and} \quad \text{(iii) } \frac{\partial F}{\partial \lambda}(x, y, \lambda) = 0.
\]

2. Solve (i) and (ii) for \(\lambda\) in terms of \(x\) and \(y\). This will give two expressions for \(\lambda\). Set them equal to each other and either solve for \(x\) in terms of \(y\) or solve for \(y\) in terms of \(x\).

3. Because \(F_\lambda = g(x, y)\), (iii) will not involve \(\lambda\). In step 2 we either solved for \(x\) in terms of \(y\) or for \(y\) in terms of \(x\); in either case, make a substitution into (iii) and solve for the remaining variable.

4. Using this known variable, substitute into (i) and (ii) to solve for the remaining two variables. (It is not always necessary to solve for \(\lambda\). This is typically just a dummy variable.)

5. Having identified all points \((x, y)\) at which an extremum might occur, check the associated values to determine which point(s) actually gives an extremum.

\(^3\)Or you can consult most standard calculus texts. The geometric reasoning underlying this strategy is fascinating.

\(^4\)Defining this function actually obscures the geometric reasoning more than I’d like, but this is in keeping with our textbook.
This is a relatively complex algorithm for optimizing a function, but we’ll try to elucidate it with some examples.

**Example 1.9.** Maximize and minimize the function $f(x, y) = xy$ subject to the constraint $4x^2 + 8y^2 = 16$.

*(Solution.)* Notice that our constraint was not given in the form $g(x, y) = 0$ for which our strategy is written. This is no problem; we simply define $g(x, y) = 4x^2 + 8y^2 - 16$ and notice that our constraint is now $g(x, y) = 0$. Having rewritten our constraint we now define

$$F(x, y, \lambda) = xy + \lambda(4x^2 + 8y^2 - 16)$$

and compute

$$F_x = y + 8\lambda x, \quad F_y = x + 16\lambda y, \quad F_\lambda = 4x^2 + 8y^2 - 16.$$ 

According to step 2 above, we now solve the equations $F_x = 0$ and $F_y = 0$ for $\lambda$:

$$8\lambda x = -y \quad \Rightarrow \quad \lambda = \frac{-y}{8x}, \quad \text{and} \quad 16\lambda y = -x \quad \Rightarrow \quad \lambda = \frac{-x}{16y}.$$ 

Still on step 2, we set these expressions equal to each other and solve for $\lambda$:

$$\frac{-y}{8x} = \frac{-x}{16y} \quad \Rightarrow \quad 16y^2 = 8x^2 \quad \Rightarrow \quad x^2 = 2y^2 \quad \Rightarrow \quad x = \pm \sqrt{2}y.$$ 

We’re now on step 3, where we substitute $x = \pm \sqrt{2}y$ into $F_\lambda = 0$. This gives

$$4(\pm \sqrt{2}y)^2 + 8y^2 - 16 = 0 \quad \Rightarrow \quad 4(2y^2) + 8y^2 = 16 \quad \Rightarrow \quad 16y^2 = 16 \quad \Rightarrow \quad y = \pm 1.$$ 

Substituting into our previous expression for $x$, we see that $x = \pm \sqrt{2}$, so we have four candidates for points which give extrema:

$$(-\sqrt{2}, -1), \quad (-\sqrt{2}, 1), \quad (\sqrt{2}, -1), \quad \text{and} \quad (\sqrt{2}, 1).$$
We see that
\[ f(-\sqrt{2}, -1) = f(\sqrt{2}, 1) = \sqrt{2} \quad \text{and} \quad f(-\sqrt{2}, 1) = f(\sqrt{2}, -1) = -\sqrt{2}, \]
so subject to our constraint, \( f \) is maximized at \((-\sqrt{2}, -1)\) and \((\sqrt{2}, 1)\) with a value of \( \sqrt{2} \) and minimized at \((-\sqrt{2}, 1)\) and \((\sqrt{2}, -1)\) with a value of \(-\sqrt{2}\). □

**Example 1.10.** In recent group work and homework we were introduced to **Cobb-Douglas production functions**, which can be used to measure production levels associated with a given amount of labor and capital. There also exist what are called **Cobb-Douglas utility functions**, which measure the level of utility (however this might be measured) that a consumer derives from a combination of goods. For example, suppose it is the case that \( x \) units of product A and \( y \) units of product B give a particular consumer utility
\[ U(x, y) = 10x^{1/4}y^{3/4}. \]
Suppose also that product A costs $40, product B costs $60, and our consumer has $800 to spend on these products. How should the consumer split this budget between the two products to maximize utility?

*(Solution.)* We have the function we’d like to maximize — our utility function — so now we need to write down our constraint in the desired form. Given the prices of the products, purchasing \( x \) units of product A and \( y \) units of product B would cost \( 40x + 60y \). Since our consumer has a budget of $800, we have the constraint
\[ 40x + 60y = 800, \]
so we let \( g(x, y) = 40x + 60y - 800 \) and define
\[ F(x, y, \lambda) = 10x^{1/4}y^{3/4} + \lambda(40x + 60y - 800). \]
We have
\[ F_x = \frac{10}{4}x^{-3/4}y^{3/4} + 40\lambda, \quad F_y = \frac{30}{4}x^{1/4}y^{-1/4} + 60\lambda, \quad \text{and} \quad F_\lambda = 40x + 60y - 800 \]
Solving \( F_x = 0 \) and \( F_y = 0 \) for \( \lambda \) then gives
\[ \frac{10}{4}x^{-3/4}y^{3/4} + 40\lambda = 0 \quad \Rightarrow \quad \lambda = \frac{-10}{160} \left( \frac{y}{x} \right)^{3/4} \]
and
\[ \frac{30}{4}x^{1/4}y^{-1/4} + 60\lambda \quad \Rightarrow \quad \lambda = \frac{-30}{240} \left( \frac{x}{y} \right)^{1/4}, \]
respectively. Setting this equal to each other, we have
\[ \frac{10}{160} \frac{y^{3/4}}{x^{3/4}} = \frac{30}{240} \frac{x^{1/4}}{y^{1/4}} \quad \Rightarrow \quad 2400y = 4800x \quad \Rightarrow \quad y = 2x. \]
Substituting this into $F_\lambda = 0$, we have

$$40x + 60(2x) - 800 = 0 \implies 40x + 120x = 800 \implies x = 800/160 = 5.$$ 

We then substitute this value into the equation $y = 2x$ to find that utility is maximized (subject to this budgetary constraint) by purchasing 5 units of product A and 10 units of product B.

We finish this section by pointing out that our strategy can easily be adapted to an optimization problem in three variables, and we illustrate this by way of example.

**Example 1.11.** Find the extreme values of $f(x, y, z) = 2x + y - 2z$ subject to the constraint $x^2 + y^2 + z^2 = 4$.

*(Solution.)* Similar to our previous strategy, we’ll start by defining an auxiliary function with an extra variable:

$$F(x, y, z, \lambda) = 2x + y - 2z + \lambda(x^2 + y^2 + z^2 - 4),$$

where $g(x, y, z) = x^2 + y^2 + z^2 - 4$ and $g(x, y, z) = 0$ is our constraint equation. We have

$$F_x = 2 + 2\lambda x, \quad F_y = 1 + 2\lambda y, \quad F_z = -2 + 2\lambda z, \quad \text{and} \quad F_\lambda = x^2 + y^2 + z^2 - 4.$$ 

We now solve $F_x = 0, F_y = 0, \text{and} F_z = 0$ for $\lambda$ to find

$$\lambda = -\frac{1}{x}, \quad \lambda = -\frac{1}{2y}, \quad \text{and} \quad F_z = \frac{1}{z}.$$ 

Setting the first two of these equal to each other gives $x = 2y$, and setting the first and last equal to each other gives $x = -z$. So $x = 2y$ and $z = -2y$. We now substitute these expressions into $F_\lambda = 0$. This gives

$$0 = (2y)^2 + y^2 + (-2y)^2 - 4 \implies 4 = 9y^2 \implies y = \pm\sqrt{4/9} = \pm\frac{2}{3}.$$ 

Since $x = 2y$ and $z = -2y$, we have two points to consider:

$$\left(\frac{4}{3}, \frac{2}{3}, -\frac{4}{3}\right) \quad \text{and} \quad \left(\frac{-4}{3}, \frac{-2}{3}, \frac{4}{3}\right).$$

Plugging into $f$ gives

$$f\left(\frac{4}{3}, \frac{2}{3}, -\frac{4}{3}\right) = \frac{18}{3} \quad \text{and} \quad f\left(\frac{-4}{3}, \frac{-2}{3}, \frac{4}{3}\right) = -\frac{18}{3},$$

so these points give a maximum and a minimum for $f$, respectively, and we see that $f$ achieves a maximum value of $18/3$ and a minimum value of $-18/3$. □
1.5 The Method of Least Squares

As a conclusion to this chapter, we’ll make reference to one of the most important applications of multivariable optimization: linear regression. In many applied settings we have some variables $x$ and $y$ which interact in a roughly (but not exactly) linear manner. For example, $x$ might be the number of students enrolled at a university and $y$ might be the number of faculty employed by that university. As the university employs more students, it will need to hire more faculty, and this relationship will be more or less linear. We could collect data over the course of a few academic terms and plot the points $(x_n, y_n)$, where $x_n$ is the student population in term $n$ and $y_n$ is the faculty population in term $n$. We then want to find an equation

$$y = mx + b$$

which roughly fits the plot. Once we have this equation, we could estimate the faculty population given the student population:

$$y_n \approx mx_n + b.$$ 

This is now a problem of choosing good values for $m$ and $b$, and we choose these values to minimize the sum of the squared errors. That is, for each term for which we have collected data, there is some error term, given by the difference between our estimate of the faculty population and the actual faculty population for this term. We square each of these errors and sum the squares and get an expression that depends on $m$ and $b$. Using our skills in optimizing two-variable functions, we can minimize this sum of squared errors term in $m$ and $b$. This strange process is illustrated by the following example.

Example 1.12. The following table gives U.S. per capita health care expenditures for the years 2005-2009:

<table>
<thead>
<tr>
<th>Year</th>
<th>Dollars</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>6,259</td>
</tr>
<tr>
<td>2006</td>
<td>7,073</td>
</tr>
<tr>
<td>2007</td>
<td>7,437</td>
</tr>
<tr>
<td>2008</td>
<td>7,720</td>
</tr>
<tr>
<td>2009</td>
<td>7,960</td>
</tr>
</tbody>
</table>

By letting $x$ be the number of years after the year 2000 and letting $y$ be the per capita health care expenditures in the given year, these points give the plot seen in Figure 12.

We would like to identify a line $y = mx + b$ which best fits this data, and will thus allow us to estimate per capita health care expenditures for later years. The graph of the line
Figure 12: Per capita health care expenditures in the U.S.

$y = mx + b$ will not pass through all five data points, but it will hopefully not miss the points by much. Notice that the error in year $n$ after 2000 is given by

$$mn + b - y_n,$$

where $y_n$ is the per capita health care expenditures in year $n$ after 2000. For example, the error in year 2008 will be $8m + b - 7,720$. We want the sum of the squares of these errors to be as small as possible\(^5\), so we would like to minimize the quantity

$$(m(5)+b-6,259)^2+(m(6)+b-7,073)^2+(m(7)+b-7,437)^2+(m(8)+b-7,720)^2+(m(9)+b-7,960)^2.$$\)

This quantity depends on our choice of $m$ and $b$ (and nothing else), so we call it $S(m,b)$. Simplifying gives

$$S(m,b) = 267,471,379 + 5b^2 - 518,384m + 255m^2 + 14b(5m - 5,207),$$

so we have

$$\frac{\partial S}{\partial m} = 510m - 518,384 + 70b \quad \text{and} \quad \frac{\partial S}{\partial b} = 10b + 14(5m - 5,207).$$

Setting $S_m = 0$ and solving for $m$ gives

$$m = \frac{259,192 - 35b}{255}.$$\)

We may substitute this into the equation $S_b = 0$ and solve for $b$ to find $b \approx 4,455.50$. We then have $m \approx 404.9$, so our line of best fit is

$$y = 404.9x + 4,455.50.$$\)

In Figure 13 we see that this line gives a good approximation to the data. We can use it to estimate that in the year 2011, per capita health care expenditures were

$$y = 404.9(11) + 4,455.50 = \$8,909.40.$$\)

---

\(^5\)If we were minimizing the sum of the errors without squaring we might have a positive error one year and a negative error the next which negate one another, but this is of course less desirable than having no error in either year. Squaring the errors avoids this confusion: the sum of squares can only be zero if our line passes through every point.
It’s exciting that we can use our calculus skills on an example such as the one above, but the process was messy and tedious — it’s not something we want to do all that often. Figure 14 provides a snippet of Mathematica code (and its output) which can be used to find the slope and $y$-intercept of our line of best fit without some of the mess.

```
(* input expenditures data. 
each entry is of the form (per capita expenditures, years since 2000) *)
healthExpenditures[x_] := Piecewise[{{6259, x == 5}, {7073, x == 6}, {7437, x == 7}, {7729, x == 8}, {7960, x == 9}}];
(* our approximation of the health expenditures depends on x, the number of years since 2000. *)
approxExpend[x_, m_, b_] := m*x + b;
(* the sum of squared errors will depend on our choices of m and b. *)
squaredErrorSum[m_, b_] := Sum[(healthExpenditures[x] - approxExpend[x, m, b])^2, {x, 5, 9}];
(* take the m-derivative and b-derivative of squaredErrorSum, and set these to 0. solve for m and b. *)
Solve[D[squaredErrorSum[m, b], m] = 0 \&\& D[squaredErrorSum[m, b], b] = 0, {m, b}] // N
```

Figure 14: Mathematica code for finding a line of best fit.

What’s more, the process we followed in this example can be carried out in more general terms to obtain the following formulae:

**Line of best fit.** Suppose we have data points $(x_1, y_1), \ldots, (x_N, y_N)$. Then the least squared-error line of best fit is given by

$$y = mx + b,$$

where

$$m = \frac{N \sum_{i=1}^{N} x_i y_i - (\sum_{i=1}^{N} x_i)(\sum_{i=1}^{N} y_i)}{N \sum_{i=1}^{N} x_i^2 - (\sum_{i=1}^{N} x_i)^2},$$

and

$$b = \frac{\sum_{i=1}^{N} y_i - m \sum_{i=1}^{N} x_i}{N}.$$
Example 1.13. In Example 1.12 we have $N = 5$ and

\[
\sum_{i=1}^{N} x_i y_i = (5)(6,259) + (6)(7,073) + (7)(7,437) + (8)(7,720) + (9)(7,960) = 259,192,
\]

\[
\sum_{i=1}^{N} x_i = 5 + 6 + 7 + 8 + 9 = 35,
\]

\[
\sum_{i=1}^{N} y_i = 6,259 + 7,073 + 7,437 + 7,720 + 7,960 = 36,449,
\]

\[
\sum_{i=1}^{N} x_i^2 = 5^2 + 6^2 + 7^2 + 8^2 + 9^2 = 255.
\]

This means that

\[
m = \frac{5(259,192) - (35)(36,449)}{5(255) - (35)^2} \approx 404.9,
\]

so

\[
b = \frac{36,449 - (404.9)(35)}{5} \approx 4,455.5,
\]

exactly as we found above.
2 Matrices

In this chapter we draw on discussions and examples found in [1] and [3].

2.1 Systems of Linear Equations with Unique Solutions

We’ll motivate this section by starting off with an example.

Example 2.1. The Tyler Table Company manufactures wooden tables and chairs. Each chair they manufacture requires 2 hours of labor and $20 worth of materials, while each table requires 4 hours of labor and $50 worth of materials. The company employs 6 workers at 40 hours/week, and these labor costs are fixed — regardless of the amount of labor needed for tables and chairs each week. The company also has $2,600 worth of materials which (for some strange reason) must be used next week or discarded. How many tables and chairs should the Tyler Table Company make in order to use all of the labor and materials they have paid for?

(Solution.) This is an admittedly contrived example, but it’s not too far off from the types of decisions many companies make in reality. The company wants to make precisely the right combination of tables and chairs to use 240 labor-hours (6 workers at 40 hours each) and to use $2,600 worth of materials, so first let’s write down the amount of labor and capital the company will need in order to make \( x \) chairs and \( y \) tables. At 2 labor-hours per chair and 4 labor-hours per table, the company will need \( 2x + 4y \) labor-hours. Since we want to use exactly 240 labor-hours, this means that we want \( x \) and \( y \) to satisfy

\[ 2x + 4y = 240. \]  

(5)

On the other hand, since each chair requires $20 worth of materials and each table requires $50 worth of materials, the total material cost for producing \( x \) chairs and \( y \) tables is \( 20x + 50y \). So \( x \) and \( y \) must also satisfy

\[ 20x + 50y = 2600. \]  

(6)

We now need to find any point(s) \( (x, y) \) that satisfy (5) and (6) simultaneously, if any such points exist. Notice that if \( (x, y) \) satisfies (5), then \( 2x = 240 - 4y \), so \( x = 120 - 2y \). We can then drop this expression for \( x \) into (6) to see that \( y \) must satisfy

\[ 20(120 - 2y) + 50y = 2600 \quad \Rightarrow \quad 2400 - 40y + 50y = 2600 \quad \Rightarrow \quad 10y = 200. \]

So \( y = 20 \). Since \( x = 120 - 2y \), we see that \( x = 80 \). So there is only one combination of chairs and tables that will allow the Tyler Table Company to exhaust its labor and capital: the company must manufacture 80 chairs and 20 tables.

Equations (5) and (6) in the above example together form what is called a system of linear equations, and such systems are very common in applications. We very frequently need to choose some combination of products or other items in order to satisfy a particular...
set of constraints. When these constraints are linear, as they are in Example 2.1, we obtain a system of linear equations. We call a point \((x, y)\) that satisfies each of the equations simultaneously a **solution to the system of linear equations**. Now the system we have in Example 2.1 is a **system of linear equations in two variables**; we will also consider systems in three variables, where the solutions have the form \((x, y, z)\), and occasionally systems in four variables, where the solutions have the form \((x, y, z, w)\). (Systems of linear equations can have as many variables as they like, but we won’t consider more than four variables.) Because the system of linear equations in Example 2.1 had only two variables, our somewhat ad hoc substitution method worked well. When we consider a greater number of variables, however, this approach becomes much less tenable, as the following example illustrates.

**Example 2.2.** Find a solution \((x, y, z)\) to the following system of linear equations, if a solution exists:

\[
\begin{align*}
  x - 2y + z &= 0 \\
  2y - 8z &= 8 \\
  5y - 4x + 9z &= -9.
\end{align*}
\]  

(Solution.) For the sake of demonstrating its ineffectiveness, we’ll use a substitution method, as in Example 2.1. First we can rewrite the second equation of our system to see that \(2y = 8z + 8\). Dividing by 2 gives \(y = 4z + 4\). Next, we substitute this expression for \(y\) into the first equation in the system to obtain

\[
  x - 2(4z + 4) + z = 0 \quad \Rightarrow \quad x - 8z - 8 + z = 0 \quad \Rightarrow \quad x - 7z - 8 = 0.
\]

This last expression can be rewritten as \(x = 7z + 8\). Finally, we substitute \(x = 7z + 8\) and \(y = 4z + 4\) into the third equation in our system. This gives

\[
  5(4z + 4) - 4(7z + 8) + 9z = -9 \quad \Rightarrow \quad 20z + 20 - 28z - 32 + 9z = -9 \quad \Rightarrow \quad z = -9.
\]

From the last expression we see that \(z = 3\); plugging this value into our earlier expressions for \(x\) and \(y\), we find that \(x = 29\) and \(y = 16\). So (7) has one solution: the point \((29, 16, 3)\).

□

Our substitution method got the job done, but it was certainly more tedious than the two-variable case. Since the systems of linear equations that show up in applications often have even more variables than three, it would be nice if we had a more systematic approach to finding solutions — something other than just following our noses, as we’ve done in these first two examples. Thankfully, such a systematic approach does exist, and it’s called the **Gauss-Jordan elimination method**. Before codifying the method explicitly, we’ll rework Example 2.1 using this method. This method will actually make Example 2.1 more tedious than it was with our substitution method, but the point here is just to show how the method works. Some of the important steps will be bold for emphasis.

**Example 2.3.** Solve the system of equations given by equations (5) and (6).
(Solution.) First, recall the system:

\begin{align*}
20x + 50y &= 2600 \\
2x + 4y &= 240.
\end{align*} \tag{8}

Notice that we’ve **interchanged the order of the equations**; certainly this will not affect the solution. Next, notice that each of the numbers in row 1 end with a zero; this equation would at least appear simpler if we divide each number by 10. It is always safe to **replace an equation with a nonzero multiple of itself**, so in this case we’ll replace row 1 with row 1 divided by 10:

\begin{align*}
2x + 5y &= 260 \\
2x + 4y &= 240.
\end{align*} \tag{9}

Similarly, we may **multiply row 2 by -1**:

\begin{align*}
2x + 5y &= 260 \\
-2x - 4y &= -240.
\end{align*} \tag{10}

Next, since each of these equations is true for our solution \((x, y)\), their sum should be true as well. That is, \((x, y)\) should be a solution to the equation given by

\[
\frac{2x + 5y = 260}{+(-2x - 4y = -240)}
\]

\[0x + y = 20.\]

Since \((x, y)\) must satisfy this new equation (notice that the equation tells us what \(y\) must be), we may **replace row 2 with the sum of row 1 and row 2**: to obtain a new, equivalent system of equations:

\begin{align*}
2x + 5y &= 260 \\
0x + y &= 20.
\end{align*} \tag{11}

Our second-to-last step is even more complicated: since we can multiply rows by nonzero numbers and replace rows with differences of the existing rows, we can **replace row 1 with row 1 minus 5 times row 2**. That is we’re going to replace row 1 with

\[
\frac{2x + 5y = 260}{-5(0x + y = 20)}
\]

\[2x + 0y = 160.\]

So we now have the following system of linear equations:

\begin{align*}
2x + 0y &= 160 \\
0x + y &= 20.
\end{align*} \tag{12}

Dividing row 1 by 2 gives us

\begin{align*}
x + 0y &= 80 \\
0x + y &= 20,
\end{align*} \tag{13}

and we can now read off our solution: \(x = 80\) and \(y = 20\). \qed
This approach to the system of linear equations in Example 2.1 was much less direct than our original approach, but we should be able to generalize the strategy used here for use on more complex problems. Each of the bold steps above were instances where we made use of the following elementary row operations, which are always allowed when manipulating systems of linear equations.

<table>
<thead>
<tr>
<th>Elementary Row Operations. When manipulating a system of linear equations, we may apply any of the following operations:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Interchange any two equations.</td>
</tr>
<tr>
<td>2. Multiply (or divide) any equation by a nonzero number.</td>
</tr>
<tr>
<td>3. Change an equation by adding to it (or subtracting from it) a multiple of another equation.</td>
</tr>
</tbody>
</table>

We have a bit of shorthand that we like to associate with each of these rules. For example, if we use operation 1 to interchange rows 2 and 3 in some system, we’ll often denote this step as \( R_2 \leftrightarrow R_3 \). If we want to multiply row 1 by 3, we’ll write \( 3R_1 \). And if we are replacing row 1 with row 1 minus 5 times row 2 (as we did in the above example), we’ll write \( R_1 - 5R_2 \). Starting with the system given to us by equations (5) and (6), here are the steps we used to solve the above system of equations, with each of the systems 8 through 13 listed:

\[
\begin{align*}
\{ \begin{array}{l}
2x + 4y &= 240 \\
20x + 50y &= 2600
\end{array} \} & \xrightarrow{R_1 \leftrightarrow R_2} & \begin{array}{l}
20x + 50y &= 2600 \\
2x + 4y &= 240
\end{array} & \xrightarrow{\frac{1}{10}R_1} & \begin{array}{l}
2x + 5y &= 260 \\
2x + 4y &= 240
\end{array} \\
\{ \begin{array}{l}
2x + 5y &= 260 \\
-20x - 250y &= -2600
\end{array} \} & \xrightarrow{-R_2} & \begin{array}{l}
2x + 5y &= 260 \\
-2x - 4y &= -240
\end{array} & \xrightarrow{R_2 + R_1} & \begin{array}{l}
2x + 5y &= 260 \\
0x + y &= 20
\end{array} \\
\{ \begin{array}{l}
2x + 0y &= 160 \\
0x + y &= 20
\end{array} \} & \xrightarrow{R_1 - 5R_2} & \begin{array}{l}
2x + 0y &= 160 \\
0x + y &= 20
\end{array} & \xrightarrow{1/2R_1} & \begin{array}{l}
x + 0y &= 80 \\
0x + y &= 20
\end{array}.
\end{align*}
\]

These six steps give a solution to our original system of linear equations by the Gauss-Jordan elimination method. Gauss-Jordan elimination is simply the repeated application of our elementary row operations until we arrive at a system such as the last one above, from which we may read off the solution to our system (the textbook calls this a system in diagonal form, for good reason).

Before using this method to solve other systems of linear equations, we’ll introduce one more (extremely important) bit of notation: the matrix. Each row of each of the above systems follows the same recipe: some coefficient multiplied by \( x \) plus some coefficient multiplied...
by \( y \) equals some number. We can streamline the process by only writing down the relevant numbers, omitting plus and equals signs. For example, the original system corresponds to the augmented matrix
\[
\begin{bmatrix}
2 & 4 & 240 \\
20 & 50 & 2600
\end{bmatrix}
\]
The bar between the last two columns is to remind us of the location of the equals sign. Entries to the left of this bar represent coefficients, and the column to the right represents a collection of numbers by which we have “augmented” our system of equations. The Gauss-Jordan elimination method now looks like

\[
\begin{bmatrix}
2 & 4 & 240 \\
20 & 50 & 2600
\end{bmatrix}
\xrightarrow{R_1 \leftrightarrow R_2}
\begin{bmatrix}
20 & 50 & 2600 \\
2 & 4 & 240
\end{bmatrix}
\xrightarrow{\frac{1}{2}R_1}
\begin{bmatrix}
2 & 5 & 260 \\
2 & 4 & 240
\end{bmatrix}
\]
\[\rightarrow -R_2
\begin{bmatrix}
2 & 5 & 260 \\
-2 & -4 & -240
\end{bmatrix}
\xrightarrow{R_2+R_1}
\begin{bmatrix}
2 & 5 & 260 \\
0 & 1 & 20
\end{bmatrix}
\]
\[\rightarrow R_1-5R_2
\begin{bmatrix}
2 & 0 & 160 \\
0 & 1 & 20
\end{bmatrix}
\xrightarrow{\frac{1}{2}R_1}
\begin{bmatrix}
1 & 0 & 80 \\
0 & 1 & 20
\end{bmatrix}
\]

In the last matrix above, we call the 2 \( \times \) 2 matrix to the left of the bar the identity matrix in two variables. Similarly, the \( n \times n \) matrix with 1 in every diagonal entry and 0 in every other entry is called the identity matrix in \( n \) variables. The goal of the Gauss-Jordan elimination method is always to reduce the matrix to the left of the bar to the identity matrix. The solution is then given by the numbers to the right of the bar (in this case, \( x = 80 \) and \( y = 20 \)).

**Example 2.4.** Use the Gauss-Jordan elimination method (with matrix notation) to find a solution to the system of linear equations given in Example 2.2.

(*Solution.*) First we want to write the system in matrix notation. This gives
\[
\begin{bmatrix}
1 & -2 & 1 & 0 \\
0 & 2 & -8 & 8 \\
-4 & 5 & 9 & -9
\end{bmatrix}
\]

At this step it’s very important to recall that our first column corresponds to the variable \( x \), our second column to \( y \), and our third column to \( z \). In particular, the numbers in the last row of our matrix are not in the same order as the coefficients of the last equation of our system. This is because the variables in that equation are ordered \( y, x, z \) instead of \( x, y, z \). We can now apply elementary row operations to our matrix:

\[
\begin{bmatrix}
1 & -2 & 1 & 0 \\
0 & 2 & -8 & 8 \\
-4 & 5 & 9 & -9
\end{bmatrix}
\xrightarrow{R_2+4R_1}
\begin{bmatrix}
1 & -2 & 1 & 0 \\
0 & 2 & -8 & 8 \\
0 & 0 & 13 & -9
\end{bmatrix}
\xrightarrow{R_3+\frac{4}{3}R_2}
\begin{bmatrix}
1 & -2 & 1 & 0 \\
0 & 2 & -8 & 8 \\
0 & 0 & 1 & 3
\end{bmatrix}
\]
\[\rightarrow \frac{1}{2}R_2
\begin{bmatrix}
1 & -2 & 1 & 0 \\
0 & 1 & -4 & 4 \\
0 & 0 & 1 & 3
\end{bmatrix}
\xrightarrow{R_2+4R_3}
\begin{bmatrix}
1 & -2 & 1 & 0 \\
0 & 1 & 0 & 16 \\
0 & 0 & 1 & 3
\end{bmatrix}
\]
\[
\begin{align*}
R_1 - R_3 & \rightarrow \begin{bmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix} \\
R_1 + 2R_2 & \rightarrow \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}.
\end{align*}
\]

This last matrix contains the \(3 \times 3\) identity matrix, as desired. From this we can read off our solution: \(x = 29\), \(y = 16\), \(z = 3\), just as before. \(\square\)

We conclude this section with another example, where we see the entire process from application problem to solution via the Gauss-Jordan method.

**Example 2.5.** A one-pound blend of coffee uses Brazilian (50¢ per ounce), Colombian (40¢ per ounce), and Peruvian (60¢ per ounce) coffee beans and costs $8.10 per pound. The blend contains twice the weight of Brazilian beans as Colombian beans. How many ounces of each type of bean does the blend contain?

*(Solution.)* Let \(B\) represent the number of ounces of Brazilian coffee in the blend, \(C\) the number of ounces of Colombian coffee, and \(P\) the number of ounces of Peruvian coffee. Since the blend comes as a one-pound package, we must have \(B + C + P = 16\). Also, since the blend is priced at $8.10 per pound, the prices of the beans tell us that

\[50 \cdot B + 40 \cdot C + 60 \cdot P = 810.\]

That is, \(50B + 40C + 60P = 810\). Finally, since the weight of the Brazilian beans is twice that of the Colombian beans, we have \(B = 2C\). Altogether this gives the following system of linear equations:

\[
\begin{align*}
B + C + P &= 16 \\
50B + 40C + 60P &= 810 \\
B - 2C + 0P &= 0.
\end{align*}
\]

(14)

For notational ease, we rewrite our system as a matrix before applying the Gauss-Jordan elimination method:

\[
\begin{bmatrix}
1 & 1 & 1 & 16 \\
50 & 40 & 60 & 810 \\
1 & -2 & 0 & 0
\end{bmatrix}.
\]

Since our columns are ordered \(B, C, P\), the solution will be ordered \(B, C, P\) in the rightmost column when we are finished. Notice that we have a 0 in the third entry of the third row; we can only end up with a 1 in this position if we interchange this row with another, so this makes for a good first move:

\[
\begin{align*}
R_2 + R_3 & \rightarrow \begin{bmatrix} 1 & 1 & 1 & 16 \\
50 & 40 & 60 & 810 \\
1 & -3 & -1 & -16 \end{bmatrix} \\
5R_3 & \rightarrow \begin{bmatrix} 1 & 1 & 1 & 16 \\
0 & -3 & -1 & -16 \\
5 & 4 & 6 & 81 \end{bmatrix} \\
R_2 - R_1 & \rightarrow \begin{bmatrix} 0 & -3 & -1 & -16 \\
5 & 4 & 6 & 81 \end{bmatrix} \\
R_3 - 5R_1 & \rightarrow \begin{bmatrix} 1 & 1 & 1 & 16 \\
0 & -3 & -1 & -16 \\
0 & -1 & 1 & 1 \end{bmatrix}
\end{align*}
\]

\(31\)
Since we now have the $3 \times 3$ identity matrix on the left side of the bar, we have our solution. This matrix gives $B = 15/2 = 7.5$, $C = 15/4 = 3.75$, and $P = 19/4 = 4.75$, so our blend consists of 7.5 ounces of Brazilian coffee, 3.75 ounces of Colombian coffee, and 4.75 ounces of Peruvian coffee. We can now easily check that these three values satisfy our three conditions.

We’ve seen that the Gauss-Jordan elimination method gives us a systematic, easily repeated process for finding solutions to systems of linear equations. The process remains, however, a laborious, tedious task. To truly master the algorithm requires practice, so you should solve several systems of linear equations. As your skills improve you will find yourself able to perform multiple operations at once and perhaps shedding the row operation notation we’ve used here.

### 2.2 General Systems of Linear Equations

In the previous section, each system of linear equations that we considered had a solution, and indeed had a unique solution; this isn’t always the case, as we can certainly write down a system of equations which cannot be simultaneously satisfied. For example, there is no point $(x,y)$ that satisfies both of the following equations:

\[ 2x + 3y = 6 \quad \text{and} \quad 2x + 3y = 8. \]

We might have $2x + 3y = 6$ or $2x + 3y = 8$, but we certainly can’t have both, so the system of linear equations consisting of these two equations has no solution. We sometimes say that such a system is inconsistent or overdetermined. On the other hand, consider the system of equations in the variables $x, y, z$ given by

\[ x - y = 0 \quad \text{and} \quad x + y = 0. \]

These two equations require that $x$ and $y$ both be zero, but they put no restriction on $z$. So this system has infinitely many solutions — we can choose $z$ to be whatever number we like. (We sometimes call such a system underdetermined.) It turns out that there are three possible outcomes for a system of linear equations; a system can have (a) a unique solution, (b) no solution, or (c) infinitely many solutions. We can make sense of these three outcomes
Graphically, at least in two variables. Consider the following system of linear equations in two variables:

\[
\begin{align*}
Ax + By &= C \\
Dx + Ey &= F,
\end{align*}
\]

where \( A, B, C, D, E, \) and \( F \) are some real numbers. Each of the two equations in this system determines a line in the \( xy \)-plane, which we can plot, and a point \((x, y)\) satisfies the equation precisely if the line passes through this point. That is, the first line consists of all solutions to the first equation, and the second line consists of all solutions to the second equation. There are then three possibilities for how these lines might interact. They could intersect exactly one time, as in Figure 15a. In this case the point \((x, y)\) through which both lines pass is a solution to both equations, and since it’s the only point through which both lines pass, it’s the unique solution to this system. On the other hand, the lines could be parallel, as in Figure 15b. In this case the lines never intersect, so there is no solution to the system. Finally, the two equations could in fact determine the same line. This is depicted in Figure 15c, where we see that every point on the indicated line is a solution the system, meaning that the system has infinitely many solutions.

The situation is the same as we increase the number of variables (though drawing the pictures becomes more difficult): a system of linear equations has the three possible outcomes mentioned above. But how do we determine which of the three situations applies to a given linear system? In two variables we could just plot the corresponding lines, as we did above, but what about in three, four, or more variables? It turns out that our Gauss-Jordan elimination method from before will continue to work, but first we should be a bit clearer about how the process works.

In the last section we gave three elementary row operations and then said that the Gauss-Jordan elimination method was simply the repeated application of these operations. When applying these operations, however, we have a goal in mind: we want to transform our system into the identity matrix, augmented by a solution to our system. Given this concrete goal, there ought to be a systematic way of applying our operations towards this goal, and indeed
The idea is to move left-to-right, transforming each column of our matrix into its desired form, with a 1 on the diagonal and zeros elsewhere. The process of making a given entry of the matrix 1 and then “zeroing out” the remainder of the column is called pivoting.

**Pivoting.** To pivot a matrix about a given nonzero entry, (1) transform the given entry into a 1 by multiplying its row by the appropriate number, and then (2) transform all other entries in the same column into zeros.

**Example 2.6.** Pivot the following matrix about the circled entry:

\[
\begin{bmatrix}
2 & 4 & 8 \\
3 & 5 & 14
\end{bmatrix}
\]

*(Solution.)* First we transform the circled 2 into a 1, and then we “zero out” the 3 below:

\[
\begin{bmatrix}
2 & 4 & 8 \\
3 & 5 & 14
\end{bmatrix}
\xrightarrow{\frac{1}{2}R_1}
\begin{bmatrix}
1 & 2 & 4 \\
3 & 5 & 14
\end{bmatrix}
\xrightarrow{R_2-3R_1}
\begin{bmatrix}
1 & 2 & 4 \\
0 & -1 & 2
\end{bmatrix}
\]

Notice that the first column of this matrix is now in the form we would want at the end of the Gauss-Jordan elimination method, so this seems like a good first step in that process. □

The Gauss-Jordan elimination method can then be described as a sequence of pivot operations:

**The Gauss-Jordan Elimination Method.** To transform a system of linear equations into diagonal form:

1. Write the linear system as an augmented matrix.
2. Make sure that the first entry of the first column is nonzero, interchanging the first row with a row below if necessary. Then pivot about the first entry of the first row.
3. Make sure that the second entry of the second column is nonzero, interchanging the second row with a row below if necessary. Then pivot about the second entry of the second row.
4. Continue until the matrix is in the desired form, then write the system of linear equations corresponding to the matrix.

This process cannot always be carried out. At each step, we can only interchange the row we’re working on with a lower row. If neither our row nor the row we’re working on has a nonzero entry in the column we’re working on, we must move to the next column, meaning that our pivots may not all lie on the diagonal.

We often refer to this process as row reducing a matrix.

---

7In fact, it can only be taken to completion when the system has a unique solution.
Example 2.7. Find a solution to the system of linear equations given by $2x + 3y = 6$ and $3x + 2y = 6$ using the Gauss-Jordan elimination method.

(*Solution.*) First we write down the matrix corresponding to this system:

$$
\begin{bmatrix}
2 & 3 & 6 \\
3 & 2 & 6
\end{bmatrix}.
$$

We can then simplify the system by pivoting about the first entry of the first column, and then the second entry of the second column:

$$
\begin{bmatrix}
2 & 3 & 6 \\
3 & 2 & 6
\end{bmatrix} \xrightarrow{R_1} \begin{bmatrix}
1 & \frac{3}{2} & 3 \\
0 & 0 & -\frac{3}{2}
\end{bmatrix}.
$$

So we see that $(\frac{6}{5}, \frac{6}{5})$ is a solution to our system. □

Our motivation for refining the Gauss-Jordan elimination method in this section was to help us determine whether a system of linear equations has a unique solution, no solution, or infinitely many solutions. We’ve seen several times now what the Gauss-Jordan process does to systems with unique solutions — it finds the solution! Let’s see now what our process does with inconsistent systems.

Example 2.8. Apply the Guass-Jordan elimination method to the system of linear equations given by $2x + 3y = 6$ and $2x + 3y = 8$.

(*Solution.*) We encountered this system at the beginning of the current section, and it was our first example of an inconsistent system. That is, this system can’t possibly have any solutions. Nonetheless, let’s try applying our process to it:

$$
\begin{bmatrix}
2 & 3 & 6 \\
2 & 3 & 8
\end{bmatrix} \xrightarrow{R_1} \begin{bmatrix}
1 & \frac{3}{2} & 3 \\
0 & 0 & -\frac{3}{2}
\end{bmatrix} \xrightarrow{R_2} \begin{bmatrix}
1 & 0 & \frac{3}{2} \\
0 & 1 & \frac{3}{2}
\end{bmatrix}.
$$

We’re now in the situation mentioned at the end of our description of the Gauss-Jordan elimination method. The second entry of the second column is zero, so we need to do some row-switching to get a nonzero entry here. But we can’t switch with row 1, because we’ve already pivoted about the first entry of the first column. This is where we would skip to the third column if there were one. Instead, our process is finished. We can now reinterpret our matrix as a linear system:

$$
x + \frac{3}{2}y = 3 \quad \text{and} \quad 0 = 2.
$$

The system we started with was pretty clearly inconsistent; the system we ended up with is just patently ridiculous. It is often the case that we start with an inconsistent system, but the inconsistency is not obvious. The Gauss-Jordan method will bring the inconsistency to light. □
**Theorem 2.** If a system of linear equations is inconsistent, then the matrix that results from the Gauss-Jordan elimination method will include a row of the form

\[
\begin{bmatrix}
0 & \cdots & 0 & | & a
\end{bmatrix},
\]

where \(a\) is a nonzero number.

So whenever a system has an inconsistency, it will be rooted out by the Gauss-Jordan method in the form of an equation \(0 = a\), where \(a\) is nonzero. Next, let’s see how the Gauss-Jordan process handles underdetermined systems, where we have infinitely many solutions.

**Example 2.9.** Find all solutions to the linear system given by

\[
x - 3y - 5z = 0 \quad \text{and} \quad y + z = 3.
\]

*(Solution.*) Whenever we have fewer equations than variables, we know that if a solution exists, it is not unique. Let’s write the system as an augmented matrix:

\[
\begin{bmatrix}
1 & -3 & -5 & | & 0 \\
0 & 1 & 1 & | & 3
\end{bmatrix}.
\]

The first column of this matrix is already in the desired form, so we pivot about the second entry of the second column. Since this entry is already 1, we must only clear out the rest of the column:

\[
\begin{bmatrix}
1 & -3 & -5 & | & 0 \\
0 & \mathbf{1} & 1 & | & 3
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & 0 & -2 & | & 9 \\
0 & \mathbf{1} & 1 & | & 3
\end{bmatrix}.
\]

This is as far as the Gauss-Jordan process will take us, so we now interpret the resulting matrix as a system of linear equations:

\[
x - 2z = 9 \quad \text{and} \quad y + z = 3.
\]

We can then choose \(z\) to be any number we like, and once this choice is made, \(x\) and \(y\) will be determined. So our general solution is

\[
z = \text{any real number}, \quad x = 9 + 2z, \quad \text{and} \quad y = 3 - z,
\]

and since \(z\) can be any real number, this system certainly has infinitely many solutions. □

Whenever a system of linear equations has infinitely many solutions, one or more of the variables is allowed to take on any value we like, as is the case for \(z\) in Example 2.9. We call such variables **free variables**. The other variables, which depend on the free variables, are called **basic variables**. In Example 2.9, \(x\) and \(y\) are both basic variables. Whenever a column is not in **proper form** (meaning that one of its entries is 1 and all others are 0) after carrying out the Gauss-Jordan process, the associated variable is free.

36
Example 2.10. Find the general solution of the system of linear equations given by

\[
\begin{align*}
x - 7y + 6w &= 5 \\
z - 2w &= -3 \\
-x + 7y - 4z + 2w &= 7
\end{align*}
\]

Identify the basic and free variables of the solution.

(Solution.) First we write the matrix corresponding to this system:

\[
\begin{bmatrix}
1 & -7 & 0 & 6 & | & 5 \\
0 & 0 & 1 & -2 & | & -3 \\
-1 & 7 & -4 & 2 & | & 7
\end{bmatrix}
\]

We can now pivot about the first entry of the first column:

\[
\begin{bmatrix}
1 & -7 & 0 & 6 & | & 5 \\
0 & 0 & 1 & -2 & | & -3 \\
-1 & 7 & -4 & 2 & | & 7
\end{bmatrix}
\xrightarrow{R_3 + R_1}
\begin{bmatrix}
1 & -7 & 0 & 6 & | & 5 \\
0 & 0 & 1 & -2 & | & -3 \\
0 & 0 & -4 & 8 & | & 12
\end{bmatrix}
\]

Since the second and third entries of the second column are zero, we skip the second column and pivot about the second entry of the third column:

\[
\begin{bmatrix}
1 & -7 & 0 & 6 & | & 5 \\
0 & 0 & 1 & -2 & | & -3 \\
0 & 0 & -4 & 8 & | & 12
\end{bmatrix}
\xrightarrow{R_3 + 4R_2}
\begin{bmatrix}
1 & -7 & 0 & 6 & | & 5 \\
0 & 0 & 1 & -2 & | & -3 \\
0 & 0 & 0 & 0 & | & 0
\end{bmatrix}
\]

We finally move to the third entry of the fourth column. There is no way to make this entry nonzero, and no more columns to consider, so the process terminates. Our reduced system is

\[
\begin{align*}
x - 7y + 6w &= 5 \\
z - 2w &= -3
\end{align*}
\]

Because the second and fourth columns are not in proper form, \(y\) and \(w\) (the variables corresponding to these columns) are free. We then write the general solution as

\[
\begin{align*}
y &= \text{any real number} \\
w &= \text{any real number} \\
x &= 7y - 6w + 5 \\
z &= 2w - 3
\end{align*}
\]

Because \(x\) and \(z\) depend on our choice of \(y\) and \(w\), \(x\) and \(z\) are the basic variables of our general solution.
2.3 Arithmetic Operations on Matrices

So far we’ve only seen matrices used to represent systems of linear equations. They have another (perhaps more important) use as representatives of linear transformations. For example, matrices can be used to model the evolution of the electorate between election cycles. Suppose that each election cycle, 90% of voters who were members of the Republican party remain in the Republican party, while the other 10% join the Democratic party. At the same time, suppose 5% of Democrats become Republicans, while the other 95% remain in the Democratic party. If each party has 500,000 voters in one election cycle, the following matrix equation computes the number of voters each party will have in the next cycle:

\[
A \begin{bmatrix} 0.95 & 0.10 \\ 0.05 & 0.90 \end{bmatrix} \begin{bmatrix} 500,000 \\ 500,000 \end{bmatrix} = \begin{bmatrix} 525,000 \\ 475,000 \end{bmatrix}.
\] (16)

Once we learn how to interpret (16) we’ll see that in the next election there should be 525,000 Democratic voters and 475,000 Republican voters. In this equation the matrix \( A \) represents the transformation from one election cycle to the next, and this is the idea we’d like to focus on next: matrices as transformations.

Before we can really consider matrices as representing transformations, we’ll need to first make sense of expressions such as (16). To this end, we’ll introduce some new vocabulary surrounding matrices and then discuss addition and multiplication of matrices.

**Definition.** The size of a matrix is given by the number of rows and columns it has. A matrix with \( m \) rows and \( n \) columns is said to be “\( m \)-by-\( n \)”, written \( m \times n \). We occasionally call a matrix with only one row a row matrix and call a matrix with just one column a column matrix; we will call either of these types of matrices vectors. Matrices which have the same number of rows and columns are called square matrices.

**Example 2.11.** Below we have a \( 2 \times 3 \) matrix, a row matrix, a column matrix, and a square matrix, respectively:

\[
\begin{bmatrix} 2 & 7 & -3 \\ 1 & 0 & 9 \end{bmatrix}, \quad \begin{bmatrix} 0 & 9 & 1 \end{bmatrix}, \quad \begin{bmatrix} 6 \\ 2 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 & 1 \\ -2 & 2 & -4 \\ 4 & -8 & 8 \end{bmatrix}.
\]

We use lowercase letters with double subscripts to identify the entries of a matrix. For example, if the square matrix above is \( A \), then \( a_{23} = -4 \).

Now we define addition of matrices. This operation is easy, but it is important to remember that addition of matrices is only defined when the matrices involved have the same size. That is, we can’t add a \( 2 \times 3 \) matrix to a \( 5 \times 4 \) matrix, but we can add two \( 2 \times 3 \) matrices together, and we do so entry-wise. For instance,

\[
\begin{bmatrix} 2 & 7 & -3 \\ 1 & 0 & 9 \end{bmatrix} + \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 9 & 0 \\ 5 & 5 & 15 \end{bmatrix}.
\]

38
We can similarly define **subtraction** entry-wise, subject to the same restriction on size:

\[
\begin{bmatrix}
2 & 7 & -3 \\
1 & 0 & 9
\end{bmatrix} - \begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{bmatrix} = \begin{bmatrix}
1 & 5 & -6 \\
-3 & -5 & 3
\end{bmatrix}.
\]

Because scalar addition is commutative and we defined matrix addition entry-wise, **matrix addition is commutative**. That is, if \(A\) and \(B\) are matrices of the same size, then \(A + B = B + A\). We point this out because it will *not* be true for multiplication.

We’re now ready to define matrix multiplication. Unfortunately this operation is not as straightforward as addition, because we have to define matrix multiplication in a way that makes equations such as (16) work. That is, matrix multiplication is defined in a special way so that they can represent linear transformations. Just as in the case of addition, there is a size restriction on which matrices can be multiplied together. Because multiplication is more complicated than addition, its condition is more complicated as well:

**Size condition for matrix multiplication.** If \(A\) is an \(m \times n\) matrix an \(B\) is a \(k \times l\) matrix, then the product \(AB\) only exists if \(n = k\). That is, the number of columns of \(A\) must equal the number of rows of \(B\). If this condition is met, the matrix \(AB\) will have size \(m \times l\).

Notice that it will sometimes be the case that the matrix \(AB\) exists, but the matrix \(BA\) doesn’t. For example, if \(A\) is a \(2 \times 3\) matrix and \(B\) is a \(3 \times 4\) matrix, then \(AB\) is a \(2 \times 4\) matrix, but the product \(BA\) doesn’t exist at all, since \(4 \neq 2\). Since \(BA\) doesn’t exist we certainly can’t have \(AB = BA\), so we see right away that **matrix multiplication is not commutative**.

Now that we’ve identified which matrices can be multiplied together, we can start describing how these products are actually computed. For simplicity, we start with the case of a row matrix and a column matrix. Suppose \(A\) is a \(1 \times n\) row matrix and \(B\) is column matrix of size \(n \times 1\):

\[
A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.
\]

Since \(A\) has the same number of columns as \(B\) has rows, the product \(AB\) exists, and will be a \(1 \times 1\) matrix — that is, a number. This product is given by

\[
AB = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = [a_1b_1 + a_2b_2 + \cdots + a_nb_n].
\]
So the matrix product $AB$ is given by multiplying the corresponding entries of $A$ and $B$ together, and adding the resulting products, *provided that $A$ is a row matrix and $B$ is a column matrix of appropriate size.*

**Example 2.12.** Compute the following matrix products:

\[
\begin{bmatrix}
0.95 & 0.10 \\
475,000 & 475,000
\end{bmatrix}
\text{ and }
\begin{bmatrix}
0.05 & 0.90 \\
475,000 & 475,000
\end{bmatrix}.
\]

*(Solution.*) We have

\[
\begin{bmatrix}
0.95 & 0.10 \\
475,000 & 475,000
\end{bmatrix}
\begin{bmatrix}
525,000 \\
0
\end{bmatrix} = [0.95 \cdot 525,000 + 0.10 \cdot 475,000] = [498,750 + 47,500] = [546,250]
\]

and

\[
\begin{bmatrix}
0.05 & 0.90 \\
475,000 & 475,000
\end{bmatrix}
\begin{bmatrix}
525,000 \\
0
\end{bmatrix} = [0.05 \cdot 525,000 + 0.90 \cdot 475,000] = [26,250 + 427,500] = [453,750].
\]

Matrix multiplication in general can now be reduced to multiplying row matrices by column matrices:

**Matrix multiplication.** Let $A$ be an $m \times n$ matrix and let $B$ be an $n \times k$ matrix. Then $AB$ is an $m \times k$ matrix, and for all $1 \leq i \leq m$ and $1 \leq j \leq k$, the $(i,j)$-entry of $AB$ is obtained by multiplying the $i^{th}$ row of $A$ by the $j^{th}$ column of $B$. That is,

\[(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj},\]

where $(AB)_{ij}$ is the entry in the $i^{th}$ row and $j^{th}$ column of $AB$.

**Example 2.13.** Let $A$ and $B$ be the following matrices:

\[
A = \begin{bmatrix}
1 & 0 & 8 \\
4 & -3 & 2
\end{bmatrix}
\text{ and }
B = \begin{bmatrix}
0 & 2 \\
6 & -5 \\
3 & 1
\end{bmatrix}.
\]

Find their product, $AB$. (Notice that this product exists, since $A$ is $2 \times 3$ and $B$ is $3 \times 2$, and $AB$ will be a $2 \times 2$ matrix.)

*(Solution.*) To find the $(1,1)$-entry of $AB$, we’ll multiply the first row of $A$ by the first column of $B$:

\[
\begin{bmatrix}
1 & 0 & 8 \\
4 & -3 & 2
\end{bmatrix}
\begin{bmatrix}
0 & 2 \\
6 & -5 \\
3 & 1
\end{bmatrix} = \begin{bmatrix}
1 \cdot 0 + 0 \cdot 6 + 8 \cdot 3 \\
1 \cdot 2 + 0 \cdot (-5) + 8 \cdot 1
\end{bmatrix} = \begin{bmatrix}
24 \\
24
\end{bmatrix}
\]
To find the entry of $AB$ that’s in the second row and first column, we’ll multiply the second row of $A$ by the first column of $B$:

\[
\begin{bmatrix}
1 & 0 & 8 \\
4 & -3 & 2
\end{bmatrix}
\begin{bmatrix}
0 & 2 \\
6 & -5 \\
3 & 1
\end{bmatrix}
= \begin{bmatrix}
24 \\
-12
\end{bmatrix}.
\]

In a similar fashion we compute the remaining two entries of $AB$:

\[
\begin{bmatrix}
1 & 0 & 8 \\
4 & -3 & 2
\end{bmatrix}
\begin{bmatrix}
0 & 2 \\
6 & -5 \\
3 & 1
\end{bmatrix}
= \begin{bmatrix}
24 \\
-12
\end{bmatrix}.
\]

and

\[
\begin{bmatrix}
1 & 0 & 8 \\
4 & -3 & 2
\end{bmatrix}
\begin{bmatrix}
0 & 2 \\
6 & -5 \\
3 & 1
\end{bmatrix}
= \begin{bmatrix}
24 \\
-12
\end{bmatrix}.
\]

For each entry, the row of $A$ and the column of $B$ correspond to the row/column location of the product. For instance, the first row of $A$ and the second column of $B$ were multiplied together to obtain the entry of $AB$ which is in the first row and second column. □

**Example 2.14.** Use the calculations from Example 2.12 to compute the following product:

\[
\begin{bmatrix}
0.95 & 0.10 \\
0.05 & 0.90
\end{bmatrix}
\begin{bmatrix}
525,000 \\
475,000
\end{bmatrix}.
\]

This represents the second iteration of the election cycle transformation we mentioned at the beginning of this section.

*(Solution.*) We’re multiplying a $2 \times 2$ matrix by a $2 \times 1$ matrix, so the result should be a $2 \times 1$ matrix. The first entry will be given by multiplying the first row of the $2 \times 2$ matrix by the only column of the $2 \times 1$ matrix, a product we computed in Example 2.12:

\[
\begin{bmatrix}
0.95 & 0.10 \\
0.05 & 0.90
\end{bmatrix}
\begin{bmatrix}
525,000 \\
475,000
\end{bmatrix}
= \begin{bmatrix}
546,250 \\
* 
\end{bmatrix}.
\]

The second entry of the product is given by the second product we computed in Example 2.12:

\[
\begin{bmatrix}
0.95 & 0.10 \\
0.05 & 0.90
\end{bmatrix}
\begin{bmatrix}
525,000 \\
475,000
\end{bmatrix}
= \begin{bmatrix}
546,250 \\
453,750
\end{bmatrix}.
\]

So two election cycles removed from the 500,000-500,000 split between Democratic and Republican voters, there will be 546,250 Democratic voters and 453,750 Republican voters. □
Mastering matrix multiplication will require a lot of practice, so you are encouraged to compute as many products as you can.

In the world of scalar multiplication we have a very special number: 1. Any number is unchanged when multiplied by 1, so we may write

\[ x \cdot 1 = 1 \cdot x = x \]

for any number \( x \). For this reason we often call 1 the **identity** of matrix multiplication. A similar situation holds for matrix multiplication, but instead of having a single identity matrix we have a sequence of identity matrices, one corresponding to each dimension. For each integer \( n \geq 1 \), the **identity matrix** \( I_n \) is the \( n \times n \) matrix will 1 in each diagonal entry and 0 in every other entry:

\[
I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
\]

If \( A \) is an \( m \times n \) matrix, then

\[ I_mA = AI_n = A. \]

(Notice the difference in dimension depending on whether we’re left-multiplying or right-multiplying the identity matrix.) In terms of transformations, identity matrices represent transformations that do nothing. For this reason, they must be square, since multiplying by a non-square matrix results in a matrix of a different size than the one we started with.

There is one last arithmetic operation we will mention for matrices: that of **scalar multiplication**. As with addition, this operation occurs entry-wise:

| Scalar multiplication. Given a scalar (i.e., real number) \( c \) and an \( m \times n \) matrix \( A \), \( cA \) is the \( m \times n \) matrix obtained by multiplying each entry of \( A \) by \( c \). That is,

\[
c \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}.
\]

| Example 2.15. Let

\[
A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}
\]

and calculate \( 2I_2 - A^2 \).
(Solution.) First we’ll calculate $A^2$:

$$A^2 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}. $$

Since $I_2$ has 1 in the two diagonal entries and 0 in the other two entries,

$$2I_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}. $$

Finally we have

$$2I_2 - A^2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} - \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ -2 & 2 \end{bmatrix}. $$

□

In this section we introduced the notion of matrices representing linear transformations. We motivated this idea with a matrix which represents the transformation of an electorate from one election cycle to the next. But we did not address an obvious question that this presents: given a linear transformation, how do we come up with a matrix which represents this transformation? For example, how do we know that the matrix in equation (16) is the correct matrix to represent our election cycle transformation? Unfortunately this question does not lie within the scope of this course, but the curious student can consult [3] or most any introductory linear algebra textbook for an answer.

2.4 The Inverse of a Matrix

We started this chapter by using matrices to solve systems of linear equations. In Section 2.3 we claimed that matrices were in fact more useful when used to represent linear transformations. This is true, but in this section we’ll see that this role of representing linear transformations in fact subsumes the role of solving linear systems. That is, we can solve systems of linear equations even more quickly by thinking of matrices as representing linear transformations. For instance, consider the following system of linear equations:

$$2x - \frac{1}{4}y = 8 $$

$$2x + \frac{1}{4}y = 12, $$

a system we can easily solve using the skills developed in Sections 2.1 and 2.2. At the same time, consider the following matrix product:

$$\begin{bmatrix} 2 & -\frac{1}{4} \\ 2 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x - \frac{1}{4}y \\ 2x + \frac{1}{4}y \end{bmatrix}. $$

43
Since the two entries of the matrix on the right side of this equation are precisely the two equations on the left side of our system in (17), we can rewrite (17) as a matrix equation:

\[
\begin{bmatrix}
2 & -1/4 \\
2 & 1/4
\end{bmatrix}
\begin{bmatrix}
x \\
y
\end{bmatrix}
=
\begin{bmatrix}
8 \\
12
\end{bmatrix}.
\tag{18}
\]

One way to interpret this equation is that the system in (17) is actually a transformation — this transformation turns the pair \((x, y)\) into the pair \((8, 12)\), and our job is to determine the values of \(x\) and \(y\). Before discussing exactly why this is a better way to use matrices when solving systems of linear equations, let’s practice translating between linear systems and their matrix equations.

**Example 2.16.** Find the system of linear equations corresponding to the matrix equation

\[
\begin{bmatrix}
0 & 1 & 2 \\
1 & 0 & 3 \\
4 & -3 & 8
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix}
=
\begin{bmatrix}
-4 \\
-2 \\
2
\end{bmatrix}.
\tag{19}
\]

Then write the matrix equation corresponding to the linear system

\[
3x - 9y = -21 \quad \text{and} \quad 2x + 6y = 22.
\tag{20}
\]

*(Solution.)* First we can carry out the matrix multiplication on the left side of (19) to find

\[
\begin{bmatrix}
0 & 1 & 2 \\
1 & 0 & 3 \\
4 & -3 & 8
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix}
=
\begin{bmatrix}
y + 2z \\
x + 3z \\
4x - 3y + 8z
\end{bmatrix}.
\]

Since this product must equal the vector on the right side of (19), the associated system of linear equations must be

\[
y + 2z = -4 \\
x + 3z = -2 \\
4x - 3y + 8z = 2.
\]

Notice that since we wrote the variables in the same order as they appear in (19), the coefficients of our linear system correspond precisely to the entries of our matrix. We can use this observation to write the matrix equation corresponding to (20) — we simply put the coefficients of our system into a matrix and multiply this by a vector of our variables to obtain a vector containing the right side of our system:

\[
\begin{bmatrix}
3 & -9 \\
2 & 6
\end{bmatrix}
\begin{bmatrix}
x \\
y
\end{bmatrix}
=
\begin{bmatrix}
-21 \\
22
\end{bmatrix}.
\]

\[\square\]
Now that we can write systems of linear equations as matrix equations (and vice versa), we need to see what this achieves. First, consider the following equation, where $a$ and $b$ are known numbers, and $a$ is not zero:

$$ax = b.$$ 

This is an equation you’ve known how to solve for a long time; to find $x$, we simply divide both sides by $a$ and conclude that $x = b/a$. To be slightly pedantic, we could alternatively think of this process as multiplying our equation by $1/a$, and think of $1/a$ as the multiplicative inverse of $a$. That is, since $ax = b$,

$$\frac{1}{a} \cdot ax = \frac{1}{a} \cdot b,$$

so $x = \frac{b}{a}$.

We call $1/a$ the multiplicative inverse of $a$ because it is the unique real number which, when multiplied by $a$, results in the identity 1. We can solve equations such as (18) in a similar manner. First, write (18) as

$$AX = B. \tag{21}$$

Recall that (21) represents a system of linear equations, and $X$ is a vector of the variables whose values we would like to know. In the scalar version of this equation, we solve for $x$ by multiplying $b$ by the multiplicative inverse of $a$. What would this look like for our matrix equation? We would have

$$X = CB,$$

where $C$ is a matrix inverse of $A$. But what does it mean to be a matrix inverse? In the scalar version, two numbers are inverse to one another if their product is 1, the identity. Similarly, we say that two square matrices are inverse to one another if their product is the identity matrix of appropriate size.

**Definition.** Suppose $A$ is an $n \times n$ matrix. The **inverse** of $A$ is the unique matrix $A^{-1}$ so that

$$AA^{-1} = A^{-1}A = I_n,$$

if such a matrix exists.

Unfortunately not all square matrices have inverses (just like not all real numbers have inverses — what’s the inverse of 0?), but if the matrix $A$ in (21) has an inverse $A^{-1}$, we can use this matrix to determine $X$. Multiplying both sides of (21) by $A^{-1}$ gives

$$A^{-1}AX = A^{-1}B \quad \Rightarrow \quad I_nX = A^{-1}B \quad \Rightarrow \quad X = A^{-1}B.$$

In particular, we can use this strategy to solve the system of linear equations in (17). Given that

$$\begin{bmatrix} 2 & -\frac{1}{4} \\ 2 & \frac{1}{4} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \frac{1}{4} & 1/4 \\ -2 & 2 \end{bmatrix}$$

are inverse to one another, we can solve (18) by multiplying both sides by the latter matrix:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & 1/4 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} 2 & -\frac{1}{4} \\ 2 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & 1/4 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} 8 \\ 12 \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}.$$ 

So the solution to (17) is $x = 5$ and $y = 8$. 

45
Example 2.17. Given that
\[
\begin{bmatrix}
-9/2 & 7 & -3/2 \\
-2 & 4 & -1 \\
3/2 & -2 & 1/2
\end{bmatrix}
\]
is the inverse of
\[
\begin{bmatrix}
0 & 1 & 2 \\
1 & 0 & 3 \\
4 & -3 & 8
\end{bmatrix}
\]
and that
\[
\begin{bmatrix}
1/6 & 1/4 \\
-1/18 & 1/12
\end{bmatrix}
\]
is the inverse of
\[
\begin{bmatrix}
3 & -9 \\
2 & 6
\end{bmatrix}
\]
find solutions to the linear systems in Example 2.16.

(Solution.) The first system has matrix equation
\[
\begin{bmatrix}
0 & 1 & 2 \\
1 & 0 & 3 \\
4 & -3 & 8
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix} = \begin{bmatrix}
-4 \\
-2 \\
2
\end{bmatrix}.
\]
Multiplying both sides by the inverse, we find that
\[
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix} = \begin{bmatrix}
-9/2 & 7 & -3/2 \\
-2 & 4 & -1 \\
3/2 & -2 & 1/2
\end{bmatrix}^{-1} \begin{bmatrix}
-4 \\
-2 \\
2
\end{bmatrix} = \begin{bmatrix}
1 \\
2 \\
-1
\end{bmatrix}.
\]
So this system has a solution in \(x = 1, y = -2, z = -1\). The matrix equation for the second system is
\[
\begin{bmatrix}
3 & -9 \\
2 & 6
\end{bmatrix}
\begin{bmatrix}
x \\
y
\end{bmatrix} = \begin{bmatrix}
-21 \\
22
\end{bmatrix}.
\]
Again, we solve by multiplying by the inverse:
\[
\begin{bmatrix}
x \\
y
\end{bmatrix} = \begin{bmatrix}
1/6 & 1/4 \\
-1/18 & 1/12
\end{bmatrix}^{-1} \begin{bmatrix}
-21 \\
22
\end{bmatrix} = \begin{bmatrix}
2 \\
3
\end{bmatrix}.
\]
So the solution to our second system is \(x = 2, y = 3\). \(\square\)

In this section we’ve discussed how inverses of square matrices can be used, but we’ve not mentioned how they might be obtained. In the next section we’ll focus on a method for computing inverses, but for now the following formula can be used in the case that \(A\) is a 2 × 2 matrix.

**Finding the inverse of a 2 × 2 matrix.** Suppose \(A\) is a 2 × 2 matrix, given by
\[
A = \begin{bmatrix}
a & b \\
c & d
\end{bmatrix},
\]
and let \(D = ad - bc\). If \(D \neq 0\), then the inverse of \(A\) is given by
\[
A^{-1} = \frac{1}{D} \begin{bmatrix}
d & -b \\
-c & a
\end{bmatrix}.
\]
If \(D = 0\), then \(A\) does not have an inverse.
2.5 The Gauss-Jordan Method for Matrix Inversion

There are a number of methods for finding the inverse of a square matrix, but we’ll give just one, and it’s a process that will be very easy to carry out after our experience with the Gauss-Jordan elimination method.

Suppose \( A \) is an invertible \( n \times n \) matrix, and \( A^{-1} \) is its inverse, meaning that \( AA^{-1} = I_n \). Now we saw in Sections 2.1 and 2.2 that the equation \( AX = B \) can be solved by row-reducing the augmented matrix \([A|B]\) until the matrix to the left of the bar is the identity, if possible. The matrix that results on the right will be \( X \). We focused on the case where \( AX = B \) represents a system of linear equations, with \( X \) our collection of variables and \( B \) the collection of solutions, but it will also apply in this case. We’re trying to solve the equation \( AA^{-1} = I_n \) for \( A^{-1} \), so we consider the augmented matrix \([A|I_n]\). As before, we’ll row-reduce this augmented matrix until the left side is the identity matrix, if possible, and the resulting matrix on the right side of the bar will be \( A^{-1} \). We summarize this method below.

**Computing the inverse of a square matrix.** If \( A \) is an \( n \times n \) matrix, consider the augmented matrix \([A|I_n]\). If possible, apply the Gauss-Jordan elimination method to row-reduce this matrix until the left side is \( I_n \). Once this process is completed, \( A^{-1} \) is the matrix on the right side. If \( I_n \) cannot be obtained on the left by elementary row operations, then \( A \) is not invertible.

Example 2.18. In Example 2.17 we solved a system of linear equations after being given that

\[
\begin{bmatrix}
-9/2 & 7 & -3/2 \\
-2 & 4 & -1 \\
3/2 & -2 & 1/2
\end{bmatrix}
\]
is the inverse of

\[
\begin{bmatrix}
0 & 1 & 2 \\
1 & 0 & 3 \\
4 & -3 & 8
\end{bmatrix}.
\]

Verify this fact using the Gauss-Jordan algorithm for computing inverses.

*(Solution.*) We’ll row reduce the augmented matrix

\[
\begin{bmatrix}
0 & 1 & 2 & | & 1 & 0 & 0 \\
1 & 0 & 3 & | & 0 & 1 & 0 \\
4 & -3 & 8 & | & 0 & 0 & 1
\end{bmatrix}.
\]

To pivot about the \((1,1)\)-entry, we’ll need to first interchange rows 1 and 2:

\[
\begin{bmatrix}
0 & 1 & 2 & | & 1 & 0 & 0 \\
1 & 0 & 3 & | & 0 & 1 & 0 \\
4 & -3 & 8 & | & 0 & 0 & 1
\end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2}
\begin{bmatrix}
1 & 0 & 3 & | & 0 & 1 & 0 \\
0 & 1 & 2 & | & 1 & 0 & 0 \\
4 & -3 & 8 & | & 0 & 0 & 1
\end{bmatrix} \xrightarrow{R_3 - 4R_1}
\begin{bmatrix}
1 & 0 & 3 & | & 0 & 1 & 0 \\
0 & 1 & 2 & | & 1 & 0 & 0 \\
0 & -3 & -4 & | & 0 & -4 & 1
\end{bmatrix}.
\]
as desired.

Next, let’s see how this algorithm handles matrices which are not invertible.

**Example 2.19.** Using the Gauss-Jordan method for inversion, attempt to find an inverse for

\[ A = \begin{bmatrix} 2 & -3 \\ -4 & 6 \end{bmatrix}. \]

*(Solution.)* We’ll row-reduce the augmented matrix

\[ \begin{bmatrix} 2 & -3 & 1 & 0 \\ -4 & 6 & 0 & 1 \end{bmatrix}. \]

We have

\[ \begin{bmatrix} 2 & -3 & 1 & 0 \\ -4 & 6 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} \boxed{2} & -3 & 1 & 0 \\ -4 & 6 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} \boxed{2} & -3/2 & \frac{1}{2} & 0 \\ -4 & 6 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} \boxed{2} & -3/2 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \]

Now that we have a row of zeros on the left, we see that no combination of elementary row operations will ever give us the identity matrix, so \( A \) must not be invertible. Because \( A \) is \( 2 \times 2 \), we can come to the same conclusion using the criterion given in Section 2.3. For the matrix \( A \), \( D = (2)(6) - (-3)(-4) = 12 - 12 = 0 \), so \( A \) is not invertible.

In Example 2.18 we were able to successfully row-reduce our matrix to the identity matrix as part of finding its inverse. We say that any matrix which can be row-reduced to the identity is **row-equivalent** to the identity matrix, and this turns out to be equivalent to invertibility. That is, a square matrix is invertible if and only if it is row-equivalent to the identity matrix of the same size. The matrix in Example 2.19 is not row-equivalent to the identity, and thus not invertible.

**Example 2.20.** Let \( A \) be a matrix so that

\[ A \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 8 \end{bmatrix} \quad \text{and} \quad A \begin{bmatrix} -2 \\ 4 \end{bmatrix} = \begin{bmatrix} -2 \\ -16 \end{bmatrix}. \]

Find \( A \).
(Solution.) First, notice that we weren’t given the size of \( A \), but that this can be easily deduced. In both products above, \( A \) is right-multiplied by a \( 2 \times 1 \) matrix, so \( A \) must have 2 columns for this multiplication to make sense. Moreover, the result of this product is a matrix with 2 rows, so \( A \) must have 2 rows. So \( A \) is a \( 2 \times 2 \) matrix. Next, we can combine the two products above into one:

\[
A \begin{bmatrix} 3 & -2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ 8 & -16 \end{bmatrix}.
\]

Now we can multiply both sides on the right by a particular matrix inverse:

\[
A \begin{bmatrix} 3 & -2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ 2 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & -2 \\ 8 & -16 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ 2 & 4 \end{bmatrix}^{-1}.
\]

Notice that the inverse must be multiplied on the right. We now use the quick approach from last section to compute the inverse: \( D = 16 \), so

\[
\begin{bmatrix} 3 & -2 \\ 2 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} 4/16 & 2/16 \\ -2/16 & 3/16 \end{bmatrix} = \begin{bmatrix} 1/4 & 1/8 \\ -1/8 & 3/16 \end{bmatrix}.
\]

Finally,

\[
A = \begin{bmatrix} 3 & -2 \\ 8 & -16 \end{bmatrix} \begin{bmatrix} 1/4 & 1/8 \\ -1/8 & 3/16 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 4 & -2 \end{bmatrix}.
\]

One can now check that \( A \) indeed satisfies the two given equations. \( \square \)

### 2.6 Input-Output Analysis

In 1973, economist Wassily Leontief was awarded the Nobel Prize in Economics, largely for his pioneering work on input-output analysis, a way of studying the interactions of various sectors of an economy. His work made great use of matrix algebra, and as a conclusion to this chapter we’ll give a brief sketch of the ideas of input-output analysis.

Our primary example in this section is reproduced from [3]. In it, we consider an economy that consists of three sectors: manufacturing, agriculture, and the service industry. As can be seen in Figure 17, each of these sectors delivers goods (or services) to consumers, but each sector also requires goods from the other sectors. For example, to produce $1 worth of goods, the manufacturing sector might have to spend 50¢ on other manufacturing goods, 20¢ on agricultural goods, and 10¢ on services. The manufacturing sector’s consumption vector would then be

\[
\begin{bmatrix} 0.5 \\ 0.2 \\ 0.1 \end{bmatrix}.
\]

We can then determine the manufacturing sector’s demand (that is, the goods it demands from other sectors) for a given output. For instance, if the manufacturing sector produces
Figure 16: Nobel laureate Wassily Leontief.

$800$ million worth of goods, we compute

\[
\begin{bmatrix}
800 \\
0.5 \\
0.2 \\
0.1
\end{bmatrix} \begin{bmatrix}
400 \\
160 \\
80
\end{bmatrix} = \begin{bmatrix}
400 \\
160 \\
80
\end{bmatrix}.
\]

(Our units above are millions of dollars.) This means that to produce $800$ million worth of goods, the manufacturing industry will have to consume $400$ million worth of manufacturing goods, $160$ million worth of goods from agriculture, and $80$ million worth of services.

So the economy has demand coming from two sources: consumers and itself. That is, while the three industries try to meet consumer demand, they generate more demand for themselves and for the other industries, thus increasing total demand. We call this demand that is generated by the economy intermediate demand, so that we have

\[
\text{total demand} = \text{intermediate demand} + \text{consumer demand}.
\]

This leads to the question that interested Leontief:

**The question of input-output analysis.** Given a certain level of consumer demand for each industry, how much should each industry produce in order to sustain itself and meet this demand?

Towards answering this question, let’s write down our assumptions about the needs of the agriculture and services industries.

<table>
<thead>
<tr>
<th>Input needed for $1$ of output</th>
<th>Manufacturing</th>
<th>Agriculture</th>
<th>Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>50¢</td>
<td>40¢</td>
<td>20¢</td>
</tr>
<tr>
<td>Agriculture</td>
<td>20¢</td>
<td>30¢</td>
<td>10¢</td>
</tr>
<tr>
<td>Services</td>
<td>10¢</td>
<td>10¢</td>
<td>30¢</td>
</tr>
</tbody>
</table>
Each column represents the demand generated by producing $1$ of output in the given sector. For example, generating $1$ of output in the services industry requires $20\text{¢}$ of input from manufacturing, $10\text{¢}$ of input from agriculture, and $30\text{¢}$ from services. We saw through the example of the manufacturing industry that each of these columns is a consumption vector, and can be used to calculate the needs of each industry for a given production level. Call these columns $c_1, c_2,$ and $c_3,$ respectively. If we want our economy to produce $x_1$ dollars worth of manufacturing, $x_2$ dollars worth of agriculture, and $x_3$ dollars worth of services, then our intermediate demand vector will be given by

$$x_1c_1 + x_2c_2 + x_3c_3.$$ 

That is, the above sum will give a vector with three entries, representing the economy’s demand for manufacturing, agriculture, and services, respectively, and not including consumers’ demand. We can write the intermediate demand vector more concisely by letting $C$ be the consumption matrix — the matrix whose columns are the columns of our above table:

$$C = \begin{bmatrix} 0.5 & 0.4 & 0.2 \\ 0.2 & 0.3 & 0.1 \\ 0.1 & 0.1 & 0.3 \end{bmatrix}. \quad (22)$$

Now our intermediate demand vector is given by

$$C \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = Cx,$$

where $x$ is what we call our production vector, representing the production levels of the three sectors. Our goal is for the production vector to match the total demand vector, which is the sum of the consumer and intermediate demand vectors.

**The Leontief Input-Output Production Equation.** For a given economy, let $C$ be the consumption matrix of this economy and let $d$ be the consumer demand vector. Then $x$, the production vector that satisfies total demand exactly, is given by

$$x = Cx + d. \quad (23)$$
This equation can be interpreted as “total production equals intermediate demand plus consumer demand”. If our economy is divided into \( n \) sectors, then \( x \) and \( d \) are \( n \times 1 \) matrices, while \( C \) is an \( n \times n \) matrix. Recalling that \( x = I_n x \), we can rewrite (23) as \( I_n x = C x + d \). Then

\[
I_n x - C x = d \quad \Rightarrow \quad (I_n - C) x = d \quad \Rightarrow \quad x = (I_n - C)^{-1} d,
\]

assuming that the matrix \( I_n - C \) is invertible.

**Example 2.21.** Using the table given above, suppose our economy’s consumer demand is $100 million for manufacturing, $60 million for agriculture, and $40 million for services. Determine the production vector \( x \) that meets this demand.

*(Solution.)* The consumption matrix for this economy is given by (22), and we see that

\[
I_3 - C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.5 & 0.4 & 0.2 \\ 0.2 & 0.3 & 0.1 \\ 0.1 & 0.1 & 0.3 \end{bmatrix} = \begin{bmatrix} 0.5 & -0.4 & -0.2 \\ -0.2 & 0.7 & -0.1 \\ -0.1 & -0.1 & 0.7 \end{bmatrix}.
\]

We can then use the Gauss-Jordan inversion method to find that

\[
(I_3 - C)^{-1} \approx \begin{bmatrix} 2.963 & 1.852 & 1.111 \\ 0.926 & 2.037 & 0.556 \\ 0.556 & 0.556 & 1.667 \end{bmatrix}.
\]

Then

\[
x = (I_3 - C)^{-1} d \approx \begin{bmatrix} 2.963 & 1.852 & 1.111 \\ 0.926 & 2.037 & 0.556 \\ 0.556 & 0.556 & 1.667 \end{bmatrix} \begin{bmatrix} 100 \\ 60 \\ 40 \end{bmatrix} \approx \begin{bmatrix} 452 \\ 237 \\ 156 \end{bmatrix}.
\]

So to meet the consumer demand of $100 million of manufacturing, $60 million of agriculture, and $40 million of services, our economy will have to have production levels of $452 million in manufacturing, $237 million in agriculture, and $156 million in services.

In more realistic models than ours, an economy is divided into tens, even hundreds, of sectors. As a result, the computations above end up involving matrices of very large size — far too large for computation by hand. Computers can be used to compute the necessary inverses and products, but sometimes the matrices are too large for even this to be practical. Thankfully, consumption matrices have a redeeming quality: their columns always\(^8\) sum to less than 1, and this allows us to make the following approximation:

\[
(I_n - C)^{-1} \approx I_n + C + C^2 + C^3 + \cdots + C^m.
\]

By choosing a large enough number \( m \), we can use the above approximation instead of actually computing an inverse.

---

\(^8\)Why is this the case? It’s an economic fact, not a mathematical one.
3 Sets and Counting

In this chapter we draw on discussions and examples found in [1].

3.1 Sets

We’ll use this section to establish the nomenclature and notation for the rest of the chapter. Many of the concepts touched upon in this section are things we already know, but we’ll write them down here in a manner that’s probably more formal than our usual day-to-day treatment of these ideas.

First, we define a set: a set is any collection of objects. The objects that live inside of a set are called the elements of the set. For example, $W$ might be the set of teams in the American League West division of Major League Baseball. The elements of $W$ are then the Texas Rangers, the Houston Astros, the Oakland Athletics, the Seattle Mariners, and the Anaheim Angels. We often like to write a set by writing its elements inside of curly brackets:

$$W = \{\text{Rangers, Astros, Athletics, Mariners, Angels}\}.$$

Sometimes a set has too many elements to list, so instead of writing its elements inside the curly brackets, we write a description of its elements. For example, let $C$ be the set of all cities in California. Some elements of $C$ include Los Angeles, Malibu, and San Diego, but there are far too many cities in California to list on this page. Instead we write

$$C = \{\text{cities in California}\}.$$

We can then express the fact that an object is an element of a given set using the symbol “$\in$”, which we often call the element symbol. For instance:

$$\text{Rangers} \in W \quad \text{and} \quad \text{Malibu} \in C.$$

To express the fact that a particular object is not in a given set, we draw a line through the element symbol:

$$\text{Cubs} \notin W \quad \text{and} \quad \text{Houston} \notin C.$$

One important thing to note is that the order in which we list the elements of a set does not matter. For instance,

$$\{1, 2, 3, 4, 5\} \quad \text{and} \quad \{2, 4, 1, 3, 5\}$$

are the same set.

Example 3.1. Let $B$ be the set of positive integers which are less than 8, and let $D$ be the set of integers which are greater than 4 but less than 10. Write $B$ and $D$ in set notation.

---

9Okay, then what’s a collection? This is what we mean when we say that we’re just formalizing what we already know. We all know from experience what a set or collection or family of objects is, but here we’re explicitly choosing the word set as the name we’ll use for this idea.
(Solution.) We have
\[ B = \{1, 2, 3, 4, 5, 6, 7\} \quad \text{and} \quad D = \{5, 6, 7, 8, 9\}. \]
\[ \square \]

Sometimes we want to consider a set of objects which are allowed to satisfy either of a pair of conditions. For example, let \( T \) be the set of all cities in Texas, so that Houston, Fort Worth, San Antonio \( \in T \), but Malibu \( \not\in T \). Suppose \( A \) is the set of all cities in either California or Texas:
\[ A = \{\text{cities in either California or Texas}\}. \]
Then we say that \( A \) is the union of \( T \) and \( C \) — it’s the set that consists of everything which is either in \( T \) or in \( C \). On the other hand, we sometimes want to consider a set of objects that must satisfy two conditions at once. For instance, let \( E \) be the set of positive integers which are greater than 4 and less than 8:
\[ E = \{5, 6, 7\}. \]
Then every element in \( E \) is in both \( B \) and \( D \) above, and every element which is in both of these sets is in \( E \). We say that \( E \) is the intersection of \( B \) and \( D \).

**Definition.** Let \( A \) and \( B \) be sets. The union of \( A \) and \( B \), written \( A \cup B \), is the set of elements that belong to either \( A \) or \( B \) (or both). The intersection of \( A \) and \( B \), written \( A \cap B \), is the set of all elements that belong to both \( A \) and \( B \).

**Example 3.2.** Let \( A = \{1, 2, 3, *, +, &\} \), \( B = \{3, 4, +, -\} \), and \( C = \{1, 5, *, -\} \). Find the following sets:
(i) \( A \cap B \);  \ (ii) \( A \cup B \);  \ (iii) \( A \cap C \);  \ (iv) \( A \cup C \);  \ (v) \( B \cap C \);  \ (vi) \( B \cup C \).

(Solution.) We have
(i) \( A \cap B = \{3, +\} \);
(ii) \( A \cup B = \{1, 2, 3, 4, *, +, -, &\} \);
(iii) \( A \cap C = \{1, *\} \);
(iv) \( A \cup C = \{1, 2, 3, 5, *, +, -, &\} \);
(v) \( B \cap C = \{-\} \);
(vi) \( B \cup C = \{1, 3, 4, 5, *, +, -\} \).
\[ \square \]
Intersections can lead to one tricky problem: what if the two sets we’re intersecting have nothing in common? For example, if \( C \) is again the set of cities in California and \( T \) is the set of cities in Texas, there are no cities that are in both sets, so what is their intersection? For this purpose we have to introduce the empty set — the set that contains nothing. We denote the empty set by \( \emptyset \), and nothing is an element of the empty set. In our notation we have

\[
C \cap T = \emptyset.
\]

Now that we’ve described some basic operations on sets — union and intersection — we want to give a name to a relationship two sets might have. When one set is entirely contained in another, we say that the first set is a subset of the latter. For example, let \( A \) be the set consisting of the cities Malibu, Los Angeles, and Bakersfield. That is,

\[
A = \{\text{Malibu, Los Angeles, Bakersfield}\}.
\]

Since every element of \( A \) is a city in California, \( A \) is a subset of \( C \), the set of all cities in California.

**Definition.** Let \( A \) and \( B \) be sets. We say that \( A \) is a subset of \( B \), and write \( A \subseteq B \), if every element of \( A \) is also an element of \( B \).

**Example 3.3.** Consider the following sets:

\[
A = \{2, 4, 5, 6, 8\}, \quad B = \{1, 3, 5, 7\}, \quad \text{and} \quad C = \{1, 2, 3, 4, 5, 6, 7, 8\}.
\]

Determine all subset relationships between these sets.

*(Solution.)* The only subset relationships are that \( A \subseteq C \) and \( B \subseteq C \). While one of the elements of \( A \) is a subset of \( B \), not all elements of \( A \) are in \( B \), so \( A \not\subseteq B \). Notice that \( A \cap B = \{5\} \), so \( A \cap B \subseteq A \) and \( A \cap B \subseteq B \). It is always the case that the intersection of two sets is a subset of each of the original sets.

**Note.** The subset relation \( \subseteq \) is a lot like the less than or equal to relation \( \leq \). It is reflexive, meaning that every set is a subset of itself (just like \( x \leq x \) for every number \( x \)). It is also transitive: this means that if \( A \subseteq B \) and \( B \subseteq C \), then \( A \subseteq C \). Finally, it is anti-symmetric, meaning that if \( A \subseteq B \) and \( B \subseteq A \), then \( A = B \). A key difference between the subset relation and less than or equal to is that not every pair of sets is comparable. For example, the sets \( A \) and \( B \) in Example 3.3 are not comparable; \( A \) is not a subset of \( B \), but \( B \) is also not a subset of \( A \). For this reason, the subset relation is known as a partial order.

Within any particular discussion we usually restrict our attention to subsets of a particular larger set, and we often call this larger set the universal set (for this particular discussion). For example, we considered the set \( W \) earlier of baseball teams in the AL West division. In this context a good universal set \( U \) might be all 30 teams in Major League Baseball. If we are having a discussion about which states in the United States have certain
properties, a good universal set would be the set of all 50 states in the country. The primary motivation for fixing a universal set is that it allows us to define the notion of a complement: all the things other than our set.

**Definition.** Let $A$ be a set, and let $U$ be a chosen universal set (necessarily containing $A$). Then the complement of $A$, written $A'$, is defined as the set of elements in $U$ which are not in $A$.

**Example 3.4.** Let $U$ be the set of states in the United States, and let $A$ be the set of states where “campus carry” is not required — that is, states where public colleges and universities are not required to allow concealed weapons on their campuses. Find $A'$.

*(Solution.)* Since $A$ is the set of states where schools are not required to allow concealed weapons on campus, $A'$ must be the set of states where schools are required to allow these weapons. So $A'$ is the set of “campus carry states”:

$$A' = \{\text{Colorado, Idaho, Kansas, Mississippi, Oregon, Texas, Utah, Wisconsin}\}.$$  

Source: National Conference of State Legislatures.

**Example 3.5.** Let $U$ be the set of students at Pepperdine, and assume that $U$ contains each of the following sets. Let $T$ be the set of members of the tennis team, let $B$ be business majors, and let $M$ be male students.

(a) Write an expression in set notation that gives the set of female members of the tennis team.

(b) Write an expression in set notation that gives the set of male students who are not on the tennis team and not majoring in business.

(c) Describe the set $M \cap B \cap T'$ in words.

(d) Describe the set $(M' \cup B) \cap T$ in words. How is this different from $M' \cup (B \cap T)$?

*(Solution.)*

(a) $M' \cap T$.

(b) $M \cap T' \cap B'$.

(c) This is the set of male business majors who are not on the tennis team.

(d) The first set is the set of members of the tennis team who are either female or majoring in business. The second set is the set of students who are either female (but not necessarily on the tennis team) or a tennis player majoring in business. The second set contains the first, but is not the same as the first: a non-tennis playing female would be in the second set, but not the first.

□
3.2 A Fundamental Principle of Counting

One of the most basic properties of a set is its size — the number of elements the set contains. In most of mathematics this is known as the **cardinality** of the set, but in keeping with our textbook we’ll typically just call this the **size** of the set or refer to it as the number of elements in the set. We denote the number of elements in the set \( S \) by \( n(S) \). In our earlier example of the teams in the AL West, \( n(W) = 5 \). The set \( C \) of cities in California is much larger, with \( n(C) = 482 \) (source: League of California Cities). For some sets the size is not a number at all — consider the set \( \mathbb{N} \) of all nonnegative integers. This set has infinitely many elements, so \( n(\mathbb{N}) = \infty \).\(^{10}\) A problem of frequent importance is determining \( n(A) \) for a given set \( A \). This is known as a **counting** problem — a problem of counting the number of elements in a given set.

**Example 3.6.** Reagan has two jobs delivering newspapers: one for the Zushi News-Journal, and one for the Yokohama Star. To save time, he works his paper routes simultaneously. Reagan delivers the News-Journal to 34 houses each morning, and delivers the Star to 52 houses. If 14 houses take both the Zushi News-Journal and the Yokohama Star, how many houses must Reagan visit each morning?

(Solution.) Let \( Z \) be the set of houses that take the Zushi News-Journal, and let \( Y \) be the set of houses that take the Yokohama Star. From the above information, we know that \( n(Z) = 34, n(Y) = 52, \) and \( n(Z \cap Y) = 14 \), and we want to know \( n(Z \cup Y) \). In total Reagan delivers \( 34 + 52 = 86 \) newspapers, but in 14 cases he is delivering two papers to the same house. That is, 14 houses were counted twice in this sum, so we subtract 14 to conclude that Reagan visits \( 86 - 14 = 72 \) houses. In our set notation,
\[
n(Z \cup Y) = n(Z) + n(Y) - n(Z \cap Y).
\]
We subtract the intersection of the two sets to avoid double-counting houses that take both papers. \( \square \)

The technique we used in Example 3.6 of first over-counting and then subtracting the intersection is in fact a more general principle, and is something we will find quite useful in several problems.

**The Principle of Inclusion and Exclusion.** Let \( A \) and \( B \) be sets of finite size. Then
\[
n(A \cup B) = n(A) + n(B) - n(A \cap B).
\]
If \( C \) is another set of finite size, then
\[
n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(A \cap C) - n(B \cap C) + n(A \cap B \cap C).
\]

\(^{10}\)This brings us to why thinking of cardinality as “the number of elements in a set” is tricky. The set of all nonnegative integers has infinitely many elements, but the set of all nonnegative real numbers has still more. In our familiar number system there is no way to distinguish between these two sizes, but cardinal numbers can make this distinction. Indeed, the cardinalities of these sets give us two different sizes of infinity.
Example 3.7. Kristin is trying to choose a hike to do tomorrow. In her list of hikes that she might do, there are 15 within an hour’s drive from her apartment. But she’s willing to drive more than an hour to do a hike, provided the hike has either a view of the Pacific Ocean or a waterfall. On her list of hikes, 23 hikes have views of the Pacific Ocean, 5 have waterfalls, and 2 have both. Of the 5 waterfall hikes, 3 are within an hour’s drive, and of the 23 hikes
with ocean views, 9 are within an hour’s drive. Given that there is only one hike with all
three properties, how many hikes are either within an hour’s drive, have a waterfall, or have
a view of the Pacific Ocean?

(Solution.) Let \( H \) be the set of hikes within an hour’s drive, \( P \) be the set of hikes with a
view of the Pacific, and \( W \) be the set of hikes with a waterfall. We’re hoping to count the set
of hikes that has any of these three properties, which is to say the set \( H \cup P \cup W \). According
to the principle of inclusion and exclusion, the size of this set is given by

\[
\begin{align*}
    n(H \cup P \cup W) &= n(H) + n(P) + n(W) \\
    &- n(H \cap P) - n(H \cap W) - n(P \cap W) \\
    &+ n(H \cap P \cap W) \\
    &= 15 + 23 + 5 - 9 - 3 - 2 + 1 = 30.
\end{align*}
\]

So Kristin has 30 hikes either within an hour’s drive, or with views of a waterfall or the
Pacific Ocean.

We used Venn diagrams to verify the important principle of this section; while we’re here,
we’ll also use them to verify another important relationship between set operations.

**De Morgan’s Laws.** Let \( A \) and \( B \) be sets. Then

\[
(A \cup B)' = A' \cap B' \quad \text{and} \quad (A \cap B)' = A' \cup B'.
\]

In English De Morgan’s laws say that the complement of a union is the intersection of the
complements and the complement of an intersection is the union of the complements, respec-
tively. We can easily convince ourselves that these statements must be true. For instance,
suppose \( A \) and \( B \) are each circles in some larger universal set \( U \), as in Figure 19. In Figure
19a we have shaded the intersection of \( A \cap B \), and in Figure 19b we take the complement of
this intersection to obtain \( (A \cap B)' \), as on the left side of the second of De Morgan’s laws.

![Venn Diagrams](https://via.placeholder.com/150)

(a) Two intersecting sets.  
(b) The complement of the intersection.

Figure 19
In Figure 20 we have the two parties to the right hand side of the second law: $A'$ and $B'$. When we take the union of these two sets the only region that remains unshaded is $A \cap B$, as in Figure 19b. So we see that $(A \cap B)' = A' \cup B'$.

We won’t show this here (or really make much use of it in this course), but De Morgan’s Laws continue to hold as we add more sets. That is,

$$(A_1 \cup A_2 \cup \cdots \cup A_n)' = A_1' \cap A_2' \cap \cdots \cap A_n'$$

and

$$(A_1 \cap A_2 \cap \cdots \cap A_n)' = A_1' \cup A_2' \cup \cdots \cup A_n',$$

for any sets $A_1, \ldots, A_n$.

### 3.3 Venn Diagrams and Counting

As we saw in the last section, Venn diagrams are often useful for visualizing intersecting sets. They can be especially useful when we are counting the number of objects with (or without) various properties, as seen by the following example, which is taken from [1].

**Example 3.8.** The three most common colors in the 193 flags of the member nations of the United Nations are red, white, and blue, and we know the following:

- 52 flags contain all three colors;
- 103 flags contain both red and white;
- 66 flags contain both red and blue;
- 73 flags contain both white and blue;
- 145 flags contain red;
- 132 flags contain white;
• 104 flags contain blue.

Use this information to answer the following questions:

(a) How many flags contain red, but not white or blue?
(b) How many flags contain exactly one of the three colors?
(c) How many flags contain none of the three colors?
(d) How many flags contain exactly two of the three colors?
(e) How many flags contain red and white, but not blue?
(f) How many flags contain red or white, but not blue?

(Solution.) Rather than work straight through the list of questions, we’ll build a Venn diagram with its various regions labeled with their sizes, and then we will be able to answer these questions with ease.

Our first piece of information is that 52 flags contain all three colors; if $R$ is the set of flags featuring red, $W$ is the flags with white, and $B$ is the set of flags that contain blue, then this means that \(n(R \cap W \cap B) = 52\), as indicated in Figure 21. Next we know that 103 flags contain both red and white; since 52 of these flags also contain blue and have already been counted, we know that the remaining 51 contain red and white, but not blue, as seen in the pink region of our diagram. Similarly, since 66 flags contain red and blue, \(66 - 52 = 14\) contain red and blue but not white, and since 73 flags contain both white and blue, \(73 - 52 = 21\) contain white and blue but not red. We then see that there are 145 flags containing red. We have already counted $51 + 52 + 14 = 117$ of these flags, so there are 28 flags that contain red, but do not contain white or blue. We similarly find that 8 flags contain white but neither of the other colors and 17 flags contain blue but neither of the other colors. Since there are 193 nations in the UN and our Venn diagram has only 191 nations, there must be 2 nations whose flags contain none of these three colors, as indicated in the lower-right corner.

We are now ready to answer the questions by picking off answers from our diagram.

(a) How many flags contain red, but not white or blue? 28
(b) How many flags contain exactly one of the three colors? \(28 + 8 + 17 = 53\)
(c) How many flags contain none of the three colors? 2
(d) How many flags contain exactly two of the three colors? \(51 + 14 + 21 = 86\)
(e) How many flags contain red and white, but not blue? 51
(f) How many flags contain red or white, but not blue? \(28 + 14 + 8 + 21 = 71\)
3.4 The Multiplication Principle

Let’s start with an example.

**Example 3.9.** Tullio’s Italian Café has a make-your-own-pasta special, where each customer chooses a type of pasta and a type of sauce for their plate. Customers choose from four types of pasta: penne, fettuccine, spaghetti, and rigatoni, and from three sauces: tomato, alfredo, and four-cheese. How many different dishes can be had from the make-your-own-pasta special?

*(Solution.)* When selecting from the make-your-own-pasta menu, a customer has two choices to make: first a pasta, then a sauce. We can diagram these two choices as a tree:

![Diagram of pasta and sauce choices](image)
We can now read off the number of choices from the number of nodes on the right: 12. It is perhaps more instructive to notice that we have 4 choices for pasta, and then each of the pasta nodes has 3 branches for a sauce choice, thus giving us $4 \cdot 3 = 12$ total choices.

The remark made at the end of our above solution holds more generally, and it is called the **multiplication principle**.

**The Multiplication Principle.** Suppose a task requires two consecutive choices. If we have $m$ options for the first choice and $n$ options for the second choice, then we have $m \cdot n$ total options for the task.

This is a helpful principle to have, but many tasks require more than two choices. For example, consider the problem of choosing an outfit that consists of a shirt, pants, and shoes. If we must choose from 9 shirts, 4 pairs of pants, and 3 pairs of shoes, how many outfit options are available? One way to answer this question is to break our task up into sub-tasks. We can think of the process as two choices: first we choose a pant-shirt combination, and then we choose our shoes. The multiplication principle says that our total number of options is given by the product of the number of pant-shirt combinations and the number of shoes. To then compute the number of pant-shirt combinations, we simply multiply our pant options by our shirt options: $4 \cdot 9 = 36$. So we have 36 pant-shirt combination options and 3 shoe options, meaning that we have $36 \cdot 3 = 108$ total outfit options. Notice that our final number of options is given by $9 \cdot 4 \cdot 3$; that is, it is the product of the number of options we had at each stage. As before, this holds more generally:

**The Generalized Multiplication Principle.** Suppose a task requires $n$ consecutive choices. If we have $m_1$ options for the first choice, $m_2$ options for the second choice, and so on, then we have $m_1 \cdot m_2 \cdots m_n$ total options for how we complete the task.

**Example 3.10.** Within Major League Baseball, the American League has 15 teams, divided into three divisions (East, Central, and West) of 5 teams each. Each year, 5 of the 15 teams make the playoffs, and the teams are chosen as follows: the winner of each division goes to the playoffs, and 2 “Wild Card” slots are filled by the remaining teams. How many different playoff configurations could the American League have? (In this problem we do care about position: it matters whether a team makes the playoffs by winning their division or earning a Wild Card slot, and it also matters whether they are the first or second Wild Card team.)

**Solution.** We fill out our playoff bracket as a sequence of five steps: we fill in the winners of each division, then fill in the Wild Card #1 slot, then the Wild Card #2 slot. Each of our first three choices has 5 options — any team can win its division. We then come to the Wild Card #1 slot. We started with 15 teams, but 3 of these won their divisions and are not eligible for the Wild Card slots. So we have 12 options for Wild Card team #1. With
this team removed from contention for the #2 slot, we have 11 options for the Wild Card #2 slot, so there are

\[ 5 \times 5 \times 5 \times 12 \times 11 = 16,500 \]

possible configurations for the American League playoffs. The setup is the same in the National League, so all told there are

\[ 16,500 \times 16,500 = 272,250,000 \]

possible configurations for the MLB playoffs.

As a last example for this section we’ll use the generalized multiplication principle to settle an important counting question in the world of sets: given a set of a particular (finite) size, how many different subsets does this set have?

**Example 3.11.** Suppose \( A \) is a set with \( n \) elements. How many different subsets does \( A \) have?

*(Solution.)* We’ll think of this question in another way: how many ways can we build a subset of \( A \)? We can answer this question as a sequence of \( n \) choices. For each element, we decide whether or not to include this element in our subset. So we have a sequence of \( n \) choices, each of which has 2 options, meaning that we have

\[
2 \times 2 \times \ldots \times 2 = 2^n
\]

\[ n \text{ times} \]

options for how we build our subset. So \( A \) has \( 2^n \) subsets (including the empty set and \( A \) itself). We call the set containing all these sets the **power set** of \( A \).

### 3.5 Permutations and Combinations

We’ll start this section with a motivating (if silly) example.

**Example 3.12.** Suppose three people — Susie, Josie, and Todd — are hoping to be elected to the town council, which has two open seats. How many ways could we choose two of these three people to serve on the council? What changes if the two seats are Treasurer and Secretary, respectively?

*(Solution.)* In the first case — when the two seats on the council are equivalent — we are simply asking how many ways we can choose two of the three people. To answer this, we list the three ways a pair can be chosen:

\[ \{S, J\}, \quad \{S, T\}, \quad \text{and} \quad \{J, T\}. \]

The first of these means that we choose Susie and Josie to be on the council, and Todd is left out. The other two are similar. Notice that \( \{J, S\} \) is the same thing as \( \{S, J\} \); either way, Susie and Josie are on the council. When the seats on the council are named, however, there are two ways we could have Susie and Josie on the council: we could have Susie as Treasurer and Josie as Secretary, or vice versa. This gives us six total options:
Treasurer  Secretary
Susie     Josie
Josie     Susie
Susie     Todd
Todd      Susie
Josie     Todd
Todd      Josie

In both cases we are selecting two individuals from a group of three, but in the first case order does not matter, while in the second case order does matter.

This type of problem comes up a lot: we need to select $r$ items from a collection of $n$ items, and sometimes the order of our selection matters, while other times it does not. Because of how often this problem presents itself, we have a name for these two scenarios.

| Permutations and Combinations. A permutation of $n$ objects taken $r$ at a time is an arrangement of $r$ of the $n$ objects in a specific order. A combination of $n$ objects taken $r$ at a time is a selection of $r$ objects from among the $n$ objects, with order disregarded. |
| In short: when order matters, it’s a permutation; when order doesn’t matter, it’s a combination. |

In Example 3.12 we had a combination and then a permutation. When we were selecting a pair of individuals for town council, we were just interested in the number of combinations of 2 individuals from a group of 3 — order didn’t matter. Once the seats on the council had titles, however, we needed to know the number of permutations of 2 individuals from a group of 3 — order did matter.

The next thing we’d like to do is use our knowledge of the multiplication principle to count the possible permutations when selecting $r$ objects from a group of $n$ objects (say, selecting two individuals from a group of 3, or selecting 10 letters from the alphabet of 26 letters). That is, we’re going to start by focusing on the case where order matters. Choosing $r$ objects from a group of $n$ objects will require $r$ steps. We choose our first object, and we have $n$ options. When we go to choose our second object, the first object is no longer available, so we have $n - 1$ options. The multiplication principle then says that there are $n(n - 1)$ ways to choose 2 objects from a group of $n$. When we choose our third object, the neither of the first two objects is available, so we have $n - 2$ options to choose from. So there are $n(n - 1)(n - 2)$ ways to choose 3 objects from $n$. This pattern continues: there are $n(n - 1)(n - 2)(n - 3)$ ways to choose 4 objects from $n$ objects, and so on. We’re now ready to write down a general formula.
**Permutation Formula.** Let \( P(n, r) \) denote the number of permutations of \( n \) objects taken \( r \) at a time. Then

\[
P(n, r) = n(n-1)(n-2) \cdots (n-(r-1)) = n(n-1)(n-2) \cdots (n-r+1).
\]

**Example 3.13.** There are 4 seats available on a town council — treasurer, secretary, chairperson, and vice chairperson — and 7 individuals are running for these four seats. How many different ways could the four seats be filled?

(\textit{Solution.}) We’re choosing 4 individuals from a group of 7, and order does matter, so we need to know the value of \( P(7, 4) \). We have

\[
P(7, 4) = 7(7 - 1)(7 - 2)(7 - 3) = 7 \cdot 6 \cdot 5 \cdot 4 = 840,
\]

so there are 840 ways these seven individuals could fill the four seats.

One problem that may not initially sound like a permutation problem is this: how many ways can a standard deck of 52 cards be ordered? We can first treat this as a multiplication principle problem: we must choose one of 52 cards, then for our second choice we select one of the remaining 51 cards, then one of the remaining 50 cards, and so on. So there are

\[
52 \cdot 51 \cdot 50 \cdot 49 \cdots 2 \cdot 1
\]

ways to order the deck. Notice that this is the same number we get out of \( P(52, 52) \), because ordering 52 cards can also be thought of as selecting 52 cards from a deck of 52, paying attention to order. The number above is far too large for most calculators to handle (\textit{Mathematica} approximates it with a 68-digit number), but we have a notation for it: 52!, which is said aloud as “fifty-two factorial”. This problem has a more general setting:

**Ordering n elements.** The number of ways to order \( n \) objects is given by

\[
P(n, n) = n(n-1)(n-2) \cdots 2 \cdot 1 = n!,
\]

where the notation \( n! \) is said “\( n \) factorial”.

**Example 3.14.** Once the 4 members of the town council in Example 3.13 have been selected, in how many different orders could they sit at the first meeting?

(\textit{Solution.}) There are

\[
4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24
\]

different orders in which the council members could sit. If the council members are Anna, Bob, Cassie, and Dave, then they could sit in any of the following orders:

\[
ABCD \ ABDC \ ACBD \ ACDB \ ADBC \ ADDB \ BACD \ BADC
\]
\[
BCAD \ BCDA \ BDAC \ BDCA \ CBA \ CABD \ CBAD \ CBDA
\]
\[
CDAB \ CDBA \ DABC \ DACB \ DBAC \ DBCA \ DCAB \ DCBA
\]

66
We see that there are 24 different orders in which the council members could sit, verifying our above computation.

We’ve now determined the number of ways that 4 council members (with distinct positions on the council) can be chosen from a group of 7 hopefuls — 840 ways — and we’ve also determined the number of unique orders in which they could sit at the first meeting once the four have been chosen — 24 orders. Next we want to consider the number of distinct groups of 4 which can appear on the town council. What we mean by this is that each of the 24 orders above are considered the same group — in each 24 of these cases, Anna, Bob, Cassie, and Dave are all selected to the town council. If Elaine is also running for town council, there will be another 24 ways to select Anna, Bob, Cassie, and Elaine to the town council, plus 24 ways to select Anna, Bob, Elaine, and Dave, and so on. All told, we determine the number of distinct groups of four by dividing the 840 total ways to fill the seats by 24, because each group of four appears 24 times in our enumeration of the 840 options. So there are

\[
\frac{840}{24} = 35
\]

distinct groups of four that we may select from the seven town council candidates. If Fred and Genova are the last two candidates, then our 35 groups are

\[
ABCD \quad ABCE \quad ABCF \quad ABCG \quad ABDE \quad ABDF \quad ABDG \\
ABEF \quad ABEG \quad ABFG \quad ACDE \quad ACDF \quad ACDG \quad ACEF \\
ACEG \quad ACFG \quad ADEF \quad ADEG \quad ADFG \quad AEFG \quad BCDE \\
BCDF \quad BCDG \quad BCEF \quad BCEG \quad BCFG \quad BDEF \quad BDEG \\
BDFG \quad BEFG \quad CDEF \quad CDEG \quad CDFG \quad CEFG \quad DEFG
\]

For emphasis, we repeat: we compute the number of ways of selecting 4 members from 7 candidates (disregarding order) by dividing the number of ordered ways we could choose these 4 candidates, and then dividing by the number of ways we could order the 4 once chosen. Remember that when we disregard order we are taking a combination of 4 elements from the 7, so our work says that the number of combinations of 4 from a group of 7 is equal to the number of permutations of 4 from a group of 7, divided by the number of orderings of 4 elements. This holds more generally:

**Combination Formula.** Let \( C(n, r) \) denote the number of combinations of \( n \) objects taken \( r \) at a time. Then

\[
C(n, r) = \frac{P(n, r)}{r!} = \frac{n!}{r!(n-r)!}.
\]

(Our textbook shows a very strong preference for the formula in the middle above, so this is what we will tend to use.)

**Example 3.15.** Silas jogs on 5 days every week, though the particular days vary week-to-week. How many ways could he choose his 5 jogging days next week?
(Solution.) Silas is interested in selecting 5 days from among the 7 next week, with their order disregarded (Sunday, Monday, Wednesday, Friday is the same as Wednesday, Sunday, Monday, Friday), so Silas is interested in a combination of 7 elements, taken 5 at a time. According to our formula, there are
\[ C(7, 5) = \frac{P(7, 5)}{5!} = \frac{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{7 \cdot 6}{2 \cdot 1} = \frac{42}{2} = 21 \]
ways for Silas to choose his 5 jog days next week. Notice that we could have alternatively computed the number of ways for Silas to choose his 2 rest days, and we would have found
\[ C(7, 2) = \frac{P(7, 2)}{2!} = \frac{7 \cdot 6}{2 \cdot 1} = \frac{42}{2} = 21 \]
ways for Silas to make his choice.

\[ \Box \]

3.6 Further Counting Techniques

In the last section of this chapter we’ll consider a couple of counting examples that don’t fall exactly under the headings of any of our previous sections. The two examples considered here show up in many different guises, and we will consider some of these real-world applications.\textsuperscript{11}

3.6.1 Stars and Bars

Consider the following problem: we have 5 stars (which we’ll denote by *) and 3 bars (denoted by |). In how many ways can we order the stars and bars, with no two bars next to each other? That is, how many distinct sequences of 5 stars and 3 bars (such as |**|*|*| or |***|*||) exist, excluding sequences such as *||***|* which have repeated bars? Perhaps it’s tempting to think that there are 8! ways to order these 8 items, but we consider any two stars or any two bars to be indistinguishable, so that interchanging a pair of stars produces the same ordering we already had. That is, this method over-counts the number of orderings. We also must disallow orderings which feature back-to-back bars.

One way to approach the problem is this: we can write the five stars, and then ask how many choices we have in placing the three bars. Doing this, we see that there are six places we could put a bar:

\[
\Box \ast \Box \ast \Box \ast \Box \ast \Box \ast \Box
\]

From these six positions we choose three in which to place a bar, and the order of this choice doesn’t matter, since the bars are indistinguishable. The fact that we don’t have back-to-back bars means that we’re choosing positions without repetition, so there are
\[
\binom{6}{3} = \frac{6!}{3!3!} = \frac{6 \cdot 5 \cdot 4}{3 \cdot 2 \cdot 1} = 20
\]

\textsuperscript{11}The names we use for these two types of counting problems do not show up in our textbook; they are just nicknames that these problems have acquired.
ways to make this choice. More generally, suppose we have \( n \) stars and \( k \) bars. If we want to place our stars and bars so that we don’t have back-to-back bars, we must have \( k \leq n + 1 \). As long as this condition is met, we have \( n + 1 \) spots in which we can place a bar and must choose to use \( k \) of these spots, so there are

\[
\binom{n + 1}{k}
\]

sequences of stars and bars to choose from.

**Example 3.16.** Between the 2010-2011 season and the 2014-2015 season, USC and UCLA played each other 12 times in basketball, with USC managing to win just two games, and these were not back-to-back victories. In how many ways could this have happened?

*(Solution.*) If we draw a star for each UCLA victory and a bar for each USC victory, we’re asking for the number of ways we can write 10 stars and 2 bars, with no back-to-back bars. According to the above, this can be done in

\[
\binom{11}{2} = \frac{11 \cdot 10}{2 \cdot 1} = 55
\]

different ways. In reality, the sequence went

\[
|****|*********,
\]

with USC winning the first and sixth games, and UCLA winning the rest.

**Example 3.17.** A poker player has her 43 chips divided into six stacks, each of which has at least one chip. If she views the stacks left-to-right, how many different ways can she divide her chips among the six stacks?

*(Solution.*) This may not at first seem to be a stars-and-bars problem, but we can think of it as (a modified version of) one. We can think of each poker chip as a star, and between each stack we have a bar. We’re asking, then, how we can arrange 43 stars and 5 bars without having any back-to-back bars (because this would mean that a stack is empty). We also make the modification that our sequence cannot start with a bar (this would mean an empty first stack) or end with a bar (giving us an empty last stack). With these modifications, there are 42 positions in which we can place our dividing lines between stacks, and we have 5 dividing lines to draw, so there are

\[
\binom{42}{5}
\]

distinct stack sequences that the poker player may realize.
3.6.2 The MISSISSIPPI Problem

In stars-and-bars type problems, swapping the position of two stars or of two bars doesn’t change the sequence. In this section we also want to consider orderings of objects, some of which are indistinguishable from each other. Let’s start with a motivating example.

Example 3.18. Seven students from the Business Administration Division at Pepperdine have been asked to represent the school at a college fair. Three of these students are majoring in Business Administration (B), two in International Business (I), and two in Accounting (A). If we only consider their majors (and not their identities), in how many orders can the students line up for a group photo?

(Solution.) We’re asking how many orders we can make with 3 copies of the letter B, 2 copies of the letter I, and 2 copies of the letter A. One way we can answer this question is with the multiplication principle: our first step is to place the Business Administration majors, and our second step is to place the International Business majors. Once we’ve done this, the Accounting majors will fill in the last two slots. So we choose 3 of the 7 slots to fill with a B:

\[
\begin{array}{cccccc}
\_ & B & \_ & B & \_ & B & \_ \\
\end{array}
\]

and then we choose 2 of the remaining 4 slots for the letter I:

\[
\begin{array}{cccccc}
I & B & I & B & \_ & \_ \\
\end{array}
\]

The last two slots must be filled by A. At the first step we had \(C(7,3)\) ways to choose where we put the Business Administration majors, and at the second step we had \(C(4,2)\) options for our placement of the International Business majors, so there are

\[
C(7,3) \cdot C(4,2) = \frac{P(7,3)}{3!} \cdot \frac{P(4,2)}{2!} = \frac{(7 \cdot 6 \cdot 5)(4 \cdot 3)}{(3 \cdot 2 \cdot 1)(2 \cdot 1)} = 7 \cdot 6 \cdot 5 = 210
\]

orders in which the majors could arrange themselves. \(\square\)

In Example 3.18 we were asked to find the possible rearrangements of the word BBBBBIIAA, and our strategy of placing the Bs and then the Is worked: we had a two-step multiplication principle problem. But what about finding all possible rearrangements of the word MISSISSIPPI? This word has four letters, so we’ll have a three-step multiplication principle problem. If we want all possible rearrangements of the word CALIFORNIA, we’ll have a seven-step rearrangement problem, with two of our letters repeating. There must be a solution to our original example which will lead us to a more tractable strategy for answering these other questions.

Thankfully, there is such a solution. The students in our group photo could arrange themselves in 7! ways, but if the three Business Administration majors chose to rearrange themselves (leaving the International Business and Accounting majors in place), this would
not change the resulting word. Since there are $3! = 6$ ways for the Business Administration majors to rearrange themselves, we’ll divide our $7!$ orderings by 6 to account for such rearrangements. Similarly, the International Business or Accounting majors could rearrange themselves in $2! = 2$ ways each, so the total number of orderings is in fact given by

$$\frac{7!}{3!2!2!} = \frac{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{(3 \cdot 2 \cdot 1)(2 \cdot 1)(2 \cdot 1)} = \frac{840}{4} = 210,$$

the same number we found before.

**Example 3.19.** Apply the above strategy to find the number of distinct ways the letters in MISSISSIPPI can be rearranged.

*(Solution.)* The word Mississippi has 11 letters: 1 M, 4 Is, 4 Ss, and 2 Ps. We can at first arrange these letters in $11!$ ways, but we must then divide out the possible ways of rearranging a given repeated letter. All told, there are

$$\frac{11!}{4!4!2!} = \frac{11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{(4 \cdot 3 \cdot 2 \cdot 1)(4 \cdot 3 \cdot 2 \cdot 1)(2 \cdot 1)} = \frac{1,663,200}{48} = 34,650$$

distinct way of rearranging these letters.

We’ll end this section with another problem that can be treated with either strategy.

**Example 3.20.** Consider the grid in Figure 22. Assuming we do not backtrack (thus moving only down or to the right), how many paths can we make from point $A$ to point $B$? How many of these paths pass through the point $C$?

*(Solution.)* In the first case we are choosing any path from $A$ to $B$, minus those which involve backtracking. In any such path we will move down on the grid 4 times and to the right 5 times; paths are distinguished only by the order in which these moves happen. So we are spelling a 9-letter word with 4 $D$s and 5 $R$s, meaning that there are

$$\frac{9!}{4!5!} = \frac{9 \cdot 8 \cdot 7 \cdot 6}{4 \cdot 3 \cdot 2 \cdot 1} = 126$$
possible routes from point \( A \) to point \( B \), not counting those which backtrack. In the second case we are restricting ourselves to those paths which pass through the point \( C \). For this we will need the multiplication principle: first we compute the possible routes from \( A \) to \( C \), and then we multiply this by the number of possible routes from \( C \) to \( B \). To get from \( A \) to \( C \) we must make two moves downward and two moves rightward, for a total of four moves. So there are
\[
\frac{4!}{2!2!} = \frac{4 \cdot 3}{2 \cdot 1} = 6
\]
ways to get from \( A \) to \( C \). Then there are
\[
\frac{5!}{2!3!} = \frac{5 \cdot 4}{2 \cdot 1} = 10
\]
paths from \( C \) to \( B \), since such a path will involve two moves down and three moves right, so in total we have \( 6 \cdot 10 = 60 \) paths from \( A \) to \( B \) which will pass through the point \( C \). So if we choose a random path from \( A \) to \( B \), this path will take us through the point \( C \) with probability
\[
\frac{60}{126} \approx 0.4762,
\]
but this is really a matter for the next chapter. \( \square \)

### 3.6.3 Increasing Order Problems

The last type of problem we’ll consider in this section is one that sounds harder than it really is. The questions in this section are of the type: how many four-letter words can be written without repetition of letters so that the letters are in alphabetical order?

One way to solve this problem (the book’s suggestion) is to write out the English alphabet in order and then choose 22 letters to delete. This will leave us with four letters, written in order. Since there are \( C(26, 22) = 14,950 \) ways to choose the letters we’ll delete, there are 14,950 such words. Another way to solve the problem is to realize that from the 26 letters in the English alphabet, we need only choose the four letters in our word. Their order in the alphabet will determine their order in our word. So there are \( C(26, 4) = 14,950 \) such words.

**Example 3.21.** Suppose Charlotte has a list of 50 household products, no two of which have the same price. She is asked to select 10 products at random and write a short review for each of the products, posting the reviews on her website in order of increasing price. How many outcomes are possible?

(*Solution.*) We can either list out the 50 products in order of increasing price and delete 40 of them, or we can choose 10 products at random and recognize that Charlotte will put the products in order on her website. In any case, we don’t particularly care about the order when counting. We’re just counting the number of ways Charlotte might choose the 10 products, which is given by
\[
\binom{50}{40} = \binom{50}{10} = 10,272,278,170.
\]
3.6.4 Some Additional Examples

**Example 3.22.** Between 2008 and 2015, USC and UCLA played each other 8 times in football with UCLA winning 3 of those games. In how many orders could this have happened?

(*Solution.*) Let’s write a $T$ for each game the USC Trojans (unfortunately) won, and a $B$ for each game won by the UCLA Bruins. In reality the sequence of victories went

$$TTTTBBBT,$$

with USC winning the first four games, UCLA winning the next 3, and then USC winning the most recent game. Determining the sequence of victories is then a question of how many ways we can order 5 copies of the letter $T$ and 3 copies of the letter $B$. Similar to above, we can consider 8 slots (in this case, the slots are games), and we must choose 3 of these to be games won by UCLA. Once we have done this, the remaining 5 games must be those won by USC, so there are

$$C(8, 3) = \frac{P(8, 3)}{3!} = \frac{8 \cdot 7 \cdot 6}{3 \cdot 2 \cdot 1} = 56$$

ways the 3-5 record could have happened for UCLA. Notice that we could have instead chosen the 5 games won by USC, and we still would have found 56 possibilities. The quickest solution to this problem would be to recognize it as a MISSISSIPPI-style problem, with 3 uses of $B$ and 5 uses of $T$ spelling an 8-letter word, so there are

$$\frac{8!}{3!5!} = 56$$

possibilities.

**Example 3.23.** Each year the NBA Finals are played as a best-of-7 series between the champions of basketball’s Eastern and Western conferences, meaning that the first team to win 4 games wins the series (there are no ties). In the 2016 NBA Finals, the Cleveland Cavaliers defeated the Golden State Warriors in 7 games. How many ways could this have happened?

(*Solution.*) Because the Cavaliers won the series, we know that they must have won 4 of the games. They could not have won more than 4, because once a team wins 4 games, the series is over (even if 7 games have not yet been played). This means that Golden State won 3 games, so it’s tempting to ask how many ways we can order 4 Cavalier wins and 3 Warrior wins. But there’s one caveat: we know that the Cavaliers must have won the final game. If they had obtained their fourth win in, say, game 5, then the series would have been over after 5 games. So really we only need to count the number of ways the first 6 games could have gone. In these games, each team won 3 games, so we need to know how many ways we
can order 3 Cavalier wins \( (C) \) and 3 Warrior wins \( (W) \). Similar to our previous examples, we have 6 games and we must choose 3 of them to have been Warrior wins, so there are

\[
C(6, 3) = \frac{P(6, 3)}{3!} = \frac{6 \cdot 5 \cdot 4}{3 \cdot 2 \cdot 1} = 20
\]

ways the series could have gone. (In reality the sequence of victories was WWCWCCC, with the Warriors losing three consecutive games for the first time all season.)
4 Probability

In this chapter we draw on discussions and examples found in [1].

4.1 Experiments, Outcomes, Sample Spaces, and Events

As with Chapter 3, we’ll start this chapter by choosing the language with which we will
formalize many notions which are already familiar to us.

\begin{definition}
An experiment is any observable or measurable activity or occurrence,
and the possible observations or measurements from an experiment are called its outcomes. All the possible outcomes of an experiment form a set, and this set is called the sample space of the experiment. Each subset of the sample space (including the empty set and the sample space itself) is called an event.
\end{definition}

In keeping with the traditions of Chapter 3, we’ve taken ideas with which we are quite
familiar and shrouded them in plenty of jargon to make them sound fancy. Let’s keep this
up with the following examples.

\textbf{Example 4.1.} The Business Administration Division has decided to survey several of its
students, and the first question of the survey asks students their major: either Business
Administration, International Business, or Accounting. Give the sample space of this exper-
iment, along with one possible event.

\textbf{(Solution.)} This experiment — asking a student her major — has three possible outcomes:
she can reply “Business Administration”, “International Business”, or “Accounting”. The
sample space is the set of possible outcomes, so the sample space is given by

\[ S = \{ \text{Business Administration, International Business, Accounting} \} \]

One possible event is \( E = \{ \text{Business Administration, International Business} \} \), which could
be described in English as “the student majors in either Business Administration or Inter-
national Business”.

\textbf{Example 4.2.} Suppose that the second question of the survey in Example 4.1 asks students
for their gender, and the possible responses to this question are \textit{female}, \textit{male}, and \textit{prefer not to
respond}. Identify the possible outcomes of this experiment (the asking of both questions), and
give its sample space. Give (in set notation) the event “the student majors in Accounting.”

\textbf{(Solution.)} The possible outcomes of this experiment are the possible combinations \((\text{major,gender})\) that could result as answers to the two question. Since there are 3 possible responses to the first question and 3 possible responses to the second question, we know that there are 9 possible responses. That is, the sample space is a set with nine elements. In
particular, the sample space is

\[ S = \{(B, F), (B, M), (B, N), (I, F), (I, M), (I, N), (A, F), (A, M), (A, N)\}, \]
where \( B, I, \) and \( A \) stand for the three majors, respectively, and \( F, M, \) and \( N \) stand for the gender options “female”, “male”, and “prefer not to respond”, respectively. The event “the student majors in Accounting” is the subset of the sample space containing all outcomes where the student is an Accounting major: \( E = \{(A,F), (A,M), (A,N)\} \).

**Example 4.3.** We say that an event has **occurred** when one of the outcomes in this event occurs. Suppose a student answers “International Business” and “Female” to the two questions in Example 4.2. Determine which of the following events has occurred.

(a) The event \( E_1 \) given by “the student is female”.

(b) The event \( E_2 = \{(I,F), (I,M), (I,P)\} \cap \{(B,M), (I,M), (A,M)\} \).

(c) The event \( E_3 = \emptyset \).

(d) The event \( E_4 = S \).

*(Solution.)*

(a) Clearly this event has occurred, since the student is female. To be pedantic, however, we write this event as a set:

\[ E_1 = \{(B,F), (I,F), (A,F)\} \]

We then notice that \((I,F) \in E_1\), so \(E_1\) has occurred.

(b) In set notation we have

\[ E_2 = \{(I,F), (I,M), (I,P)\} \cap \{(B,M), (I,M), (A,M)\} = \{(I,M)\} \]

so \((I,F) \notin E_2\), meaning that the event \(E_2\) has not occurred. We can also conclude this from the more natural description of \(E_2\): this is the event where the student is majoring in International Business and is male.

(c) Our only real approach here is set notation. Since \((I,F) \notin \emptyset\), \(E_3\) must not have occurred. In fact, the event \(\emptyset\) can never occur, since the corresponding set has no elements, so we call this the **impossible event**.

(d) On the other hand we have \((I,F) \in S\), so \(S\) has occurred. The sample space \(S\) contains every possible outcome, so whenever an outcome has occurred, \(S\) has occurred. For this reason we call \(S\) the **certain event**.
Since the events of an experiment are sets (indeed, they are subsets of the same universal space $S$), the usual set operations exist between them. The **union** of an event $E_1$ and an event $E_2$ is the event $E_1 \cup E_2$, and is often said as “either $E_1$ or $E_2$.” Similarly, the **intersection** of $E_1$ and $E_2$ is the event corresponding to the set $E_1 \cap E_2$, and is said as “both $E_1$ and $E_2$.” Finally, the **complement** of an event $E_1$ is the complement of the set to which $E_1$ corresponds, with the universal set understood to be the sample space $S$. As with sets, the complement of $E_1$ is denoted with a prime — $E_1'$ — and the event $E_1'$ is said to have occurred precisely when the event $E_1$ has not occurred.

**Example 4.4.** Suppose we roll a pair of six-sided dice, and consider the following events:

\[ E_1 = \text{“the first roll is even”} \]
\[ E_2 = \text{“the second roll is odd”} \]
\[ E_3 = \text{“the two rolls sum to 5”} \]

Give a description of each of the following events:
(a) $E_1 \cap E_2$,  
(b) $E_1 \cup E_2$,  
(c) $(E_1 \cup E_2)'$,  
(d) $E_1 \cap E_2 \cap E_3$,  
(e) $(E_1' \cup E_2)' \cap E_3$.

(Solution.)

(a) The intersection of $E_1$ and $E_2$ is the event that occurs when both $E_1$ and $E_2$ occur, so $E_1 \cap E_2$ is the event “the first roll is even and the second roll is odd.” There are 9 ways for this to happen: we have 3 choices for our first roll (2,4,6) 3 choices for our second roll (1,3,5).

(b) This is the event “the first roll is even or the second roll is odd.” One method for counting the ways this event could happen is the inclusion/exclusion principle, since $n(E_1 \cup E_2) = n(E_1) + n(E_2) - n(E_1 \cap E_2)$. Since the first roll will be even in half of the 36 outcomes, $n(E_1) = 18$. Similarly, $n(E_2) = 18$, so $n(E_1 \cup E_2) = 18 + 18 - 9 = 27$, meaning that there are 27 ways to either have the first roll be even or the second roll be odd.

(c) This is the complement of the previous event, and we can use De Morgan’s laws to rewrite it as $E_1' \cap E_2'$. Having done this, we may say that this is the event “the first roll is not even and the second roll is not odd,” or perhaps more directly, “the first roll is odd and the second roll is even.” Just as in our first example, the multiplication principle tells us that there are 9 outcomes that lead to this event. This event also gives us another way to count the event $E_1 \cup E_2$, using the fact that these two events are complementary. We have

\[ n(E_1 \cup E_2) = n(S) - n((E_1 \cup E_2)') = 36 - 9 = 27, \]

just as before.

(d) This is the event “the first roll is even, the second roll is odd, and the two rolls sum to 5,” and there are two outcomes in this event: (2,3) and (4,1).
(e) Using De Morgan’s laws to rewrite this as $(E'_1 \cap E'_2 \cap E_3 = E_1 \cap E'_2 \cap E_3$, we see that this is the event “the first roll is even, the second roll is even, and the two rolls sum to 5,” which is to say “both rolls are even and the two rolls sum to 5.” Since there is no way for two even rolls to sum to 5, this is the impossible event, $\emptyset$.

In the above example, $E_1 \cap E'_2$ represents the event “both rolls are even,” and in the last part of the example we considered the event $(E_1 \cap E'_2) \cap E_3$. That is, the event “both rolls are even and the rolls sum to 5.” We concluded however, that this is the impossible event — there is no way for the event “both rolls are even” and the event “the two rolls sum to 5” to occur simultaneously. When two events can never occur at the same time — that is, when their intersection is the impossible event — we say that the two events are mutually exclusive.

4.2 Assignment of Probabilities

In this section we want to discuss probabilities of events. Since events are sets of outcomes, we’ll start with probabilities of outcomes. In short, the probability of an outcome is a measure of the likelihood that that outcome will occur. We measure probabilities on a scale from 0 to 1: an outcome that cannot possibly happen has probability 0, while an outcome that must happen has probability 1. For outcomes that are neither impossible nor certain, we use numbers between 0 and 1, and this number represents the frequency with which we expect the outcome to occur if the experiment is done a large number of times. For example, if we roll a die 1,000 times, we would expect to roll a 2 about a sixth of the time, so we write

$$P(2) = \frac{1}{6}.$$  

This means that the probability of rolling a 2 is $1/6 \approx 0.1667$. Because of our experience rolling dice, we can easily agree that this is the appropriate probability for rolling a 2. However, most of the examples we consider will not be so familiar to us, so let’s make our reasoning clear. In the experiment of rolling a die the sample space is

$$S = \{1, 2, 3, 4, 5, 6\}.$$  

Each of these outcomes occurs with the same likelihood\textsuperscript{(a)}, so the probability of rolling a 2 is

$$P(2) = \frac{1}{n(S)} = \frac{1}{6}.$$  

We say that we’ve computed the relative frequency of rolling a 2. There are 6 possible outcomes, and exactly 1 of these has us roll a 2, so we roll a 2 in $1/6$ of the possible outcomes. Of course, it doesn’t matter that we chose 2. The probability of rolling any particular number is $1/6$, so this experiment has probability distribution

\textsuperscript{(a)}This is an important point. Most of the time certain outcomes are more likely than others, and we must make this a part of our probability computations.
A **probability distribution** is a rule which assigns a probability to each of the outcomes in an experiment’s sample space.

**Example 4.5.** Suppose a survey is conducted asking 1,000 people for their favorite late-night talk show host, and the following table gives the results:

<table>
<thead>
<tr>
<th>Talk show host</th>
<th>No. of individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jimmy Kimmel</td>
<td>298</td>
</tr>
<tr>
<td>Jimmy Fallon</td>
<td>275</td>
</tr>
<tr>
<td>Stephen Colbert</td>
<td>199</td>
</tr>
<tr>
<td>Conan O’Brien</td>
<td>78</td>
</tr>
<tr>
<td>James Corden</td>
<td>31</td>
</tr>
<tr>
<td>Other/None</td>
<td>119</td>
</tr>
</tbody>
</table>

Suppose 1 of these 1000 individuals is selected at random and asked for her favorite talk show host. Give the probability distribution for this experiment. Find the probability that the individual’s favorite host is named “Jimmy.” Find the probability that the individual’s favorite host does not have the birth name “James.”

*(Solution.)* This question is a little tricky, because the argument could be made that our sample space is 

\[ S = \{\text{Kimmel, Fallon, Colbert, Conan, Corden, Other/None}\}. \]

But right now we only know how to work with sample spaces where each of the outcomes is equally likely, so the sample space will actually be the 1,000 possible respondents we could choose:

\[ S = \{\text{person 1, person 2, \ldots, person 1000}\}. \]

From these 1,000 possible outcomes, 298 will result in the individual telling us that her favorite late-night talk show host is Jimmy Kimmel, so the relative frequency of the answer “Jimmy Kimmel” is

\[
\frac{\text{#Jimmy Kimmel fans}}{n(S)} = \frac{298}{1,000} = 0.298.
\]

For this reason, we say that \( P(\text{Jimmy Kimmel}) = 0.298 \). We can similarly compute the other relative frequencies to find our probability distribution:

<table>
<thead>
<tr>
<th>Talk show host</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jimmy Kimmel</td>
<td>0.298</td>
</tr>
<tr>
<td>Jimmy Fallon</td>
<td>0.275</td>
</tr>
<tr>
<td>Stephen Colbert</td>
<td>0.199</td>
</tr>
<tr>
<td>Conan O’Brien</td>
<td>0.078</td>
</tr>
<tr>
<td>James Corden</td>
<td>0.031</td>
</tr>
<tr>
<td>Other/None</td>
<td>0.119</td>
</tr>
</tbody>
</table>
Now we could list our sample space as

\[ S = \{ \text{Kimmel, Fallon, Colbert, Conan, Corden, Other/None} \}, \]

recognizing that the outcomes have probabilities determined by the above table. Notice that \(298 + 275 = 573\) individuals have either Jimmy Kimmel or Jimmy Fallon as his/her favorite host, so the probability that an individual’s favorite host is named “Jimmy” is \(573/1000 = 0.573\). Including James Corden, a total of 604 fans have a favorite host with the birth name “James,” so \(396/1000 = 0.396\) gives the probability that we select an individual whose favorite host is not named James.

This example hints at two important properties of probability computations. The first is the \textbf{addition principle}. Since an event is a set of outcomes, we can compute the probability of an event by summing the probabilities of its various outcomes:

\[ \text{Addition Principle.} \quad \text{Given an event } E = \{m_1, m_2, \ldots, m_n\}, \text{ the probability of } E, \quad P(E) \text{ is given by} \]
\[ P(E) = P(m_1) + P(m_2) + \cdots + P(m_n). \]

That is, we find the probability of an event by summing the probabilities of the outcomes that make up this event.

For instance, in Example 4.5, the event “favorite talk show host is named Jimmy” consists of the outcomes “favorite talk show host is Jimmy Kimmel” and “favorite talk show host is Jimmy Fallon.” Then the probability of this event is the sum of its two outcomes:

\[ P(\text{name Jimmy}) = P(\text{Jimmy Kimmel}) = P(\text{Jimmy Fallon}) = 0.298 + 0.275 = 0.573. \]

The event \(S\) is the entire sample space — the certain event — and thus has probability 1. Since the addition principle says that the probability of an event is the sum of the probabilities of the outcomes that make up this event, we see that the sum of the probabilities of all possible events must be 1. This gives us the following properties for a probability distribution.

\[ \text{Fundamental Properties of Probability Distributions.} \quad \text{Suppose an experiment has sample space } S = \{s_1, s_2, \ldots, s_n\}, \text{ and that the probabilities corresponding to these outcomes are } p_1, p_2, \ldots, p_n. \text{ Then} \]

\begin{enumerate}
  
  \item For \(i = 1, 2, \ldots, n\), \(0 \leq p_i \leq 1\).
  \item The probabilities sum to 1: \(p_1 + p_2 + \cdots + p_n = 1\).
\end{enumerate}

In English, the second property says that one of the outcomes \(s_1, \ldots, s_n\) must occur.

The additional principle leads to a nice formula: \textit{when every outcome in a sample space has the same probability} (and only then), the probability of an event is given by its relative size in the sample space:

\[ P(E) = \frac{n(E)}{n(S)}. \]

80
We will spend more time with this formula in the next section, but for now we observe that it allows us to reformulate one of our results about sets in terms of probabilities. If we have events \( E \) and \( F \), we can think of these events as sets and recall the inclusion/exclusion principle:

\[
n(E \cup F) = n(E) + n(F) - n(E \cap F).
\]

We can then divide both sides of this equation by \( n(S) \):

\[
\frac{n(E \cup F)}{n(S)} + \frac{n(E)}{n(S)} + \frac{n(F)}{n(S)} - \frac{n(E \cap F)}{n(S)}.
\]

This gives the relative frequencies of the four events involved, so we have

\[
P(E \cup F) = P(E) + P(F) - P(E \cap F).
\]

In English this says the probability that \( E \) or \( F \) happens is the probability that \( E \) happens plus the probability that \( F \) happens, minus the probability that both happen. This is an important result, so we’ll restate it.

**Inclusion/Exclusion Principle for Probabilities.** Let \( E \) and \( F \) be events. Then

\[
P(E \cup F) = P(E) + P(F) - P(E \cap F).
\]

If \( E \) and \( F \) are mutually exclusive, then \( E \cap F = \emptyset \), so \( P(E \cap F) = 0 \) and we have

\[
P(E \cup F) = P(E) + P(F).
\]

(Warning: This second formula only applies when \( E \) and \( F \) are mutually exclusive.)

**Example 4.6.** In baseball, the World Series is an annual series of games played between the champions of the National and American leagues. The Texas Rangers play in the American League and the Los Angeles Dodgers play in the National League. As of July 17, 2016, FanGraphs said that the Rangers had a 13.8\% chance of appearing in the 2016 World Series, and that the Dodgers had a 16.4\% chance of being the National League’s representative. If the probability of these two teams facing each other in the World Series is 0.023, what is the probability that at least one of these teams makes it to the World Series?

**Solution.** We’ll let \( R \) be the event “the Rangers appear in the World Series,” and \( D \) will be the event “the Dodgers appear in the World Series.” We’ve been told that \( P(R) = 0.138 \) and \( P(D) = 0.164 \). The event \( R \cap D \) is the event that these teams are both in the World Series, so \( P(R \cap D) = 0.023 \). We can now use the inclusion/exclusion principle to conclude that

\[
P(R \cup D) = P(R) + P(D) - P(R \cap D) = 0.138 + 0.164 - 0.023 = 0.279.
\]

That is, there is a 27.9\% chance that either the Rangers or the Dodgers will make it to the World Series.
Example 4.7. The Seattle Mariners play in the American League, and as of July 17, 2016, FanGraphs gave them a 2.4% chance of making the 2016 World Series. Given the probabilities in Example 4.6, compute the probability that either the Mariners or the Rangers appear in the World Series.\(^{13}\)

(Solution.) As with Example 4.6, we want to use the principle of inclusion and exclusion, but there’s one problem: we have not been given the probability of both the Rangers and the Mariners making the World Series. This is okay, though. Because the Mariners and Rangers are both in the American League, at most one of them can represent the AL in the World Series. That is, if \(M\) is the event that the Mariners make the World Series, then \(R\) and \(M\) are mutually exclusive events. So we have

\[
P(R \cup M) = P(R) + P(M) - P(R \cap M) = 0.138 + 0.024 - 0 = 0.162,
\]

meaning that there is a 16.2% chance of either the Mariners or Rangers appearing in the 2016 World Series. \(\square\)

We’ll end this section with a quick contrast of probabilities with odds. The likelihood of an event happening is often given not by the event’s probabilities, but by its odds, a ratio between the likelihood that the event will happen and the likelihood that the event will not happen. For instance, you might be told that the odds against you winning a sweepstakes are “ten thousand to one.” This means that you are 10,000 times as likely to not win the sweepstakes as you are to win it, so the probability of not winning the sweepstakes is

\[
\frac{10,000}{10,000 + 1} = \frac{10,000}{10,001}.
\]

Unfortunately odds are variously given as “odds” and “odds against,” meaning that we must take care before interpreting any odds. For instance, when the odds against winning are 10,000 to 1, the odds of winning are 1 to 10,000.

Example 4.8. Suppose we flip a fair coin. Find the odds of flipping heads. If we flip the coin twice, find the odds against flipping heads both times.

(Solution.) When we flip a coin the probability of flipping heads is 0.5, and the probability against heads is 0.5, so the odds are

\[
0.5 : 0.5 = 1 : 1
\]

that we will flip heads. If we flip the coin twice there is a 25% chance we’ll get heads both times, and a 75% chance we will not. So the odds are

\[
0.75 : 0.25 = 3 : 1
\]

against two consecutive heads. \(\square\)

\(^{13}\)This example was originally going to feature the Anaheim Angels instead of the Seattle Mariners, but FanGraphs gave the Angels a 0% chance of making the World Series. Good!
4.3 Calculating Probabilities of Events

In this section we’ll start developing our toolbox for computing probabilities. We’ll cover an important special case of probability distributions (a case we referenced near the end of the previous section), and we’ll discuss a strategy that is often very useful in computing probabilities.

4.3.1 Experiments with Equally Likely Outcomes

In Example 4.5 we had a sample space of 1,000 individuals. Because we were randomly selecting an individual from this group, we were just as likely to select any one individual as we were any other. That is, the outcomes were equally likely. We could then compute the probability of any event by comparing the number of outcomes in this event to the total number of outcomes in the sample space. When \( E \) was the event “the individual’s favorite host is Jimmy Kimmel,” we found that

\[
P(E) = \frac{n(E)}{n(S)} = \frac{298}{1,000} = 0.298.
\]

This strategy works whenever the various outcomes in a sample space are all equally likely.

**Relative Frequency.** Let \( S \) be a sample space (of finite size) whose outcomes are all equally likely. If \( E \) is any event, then

\[
P(E) = \frac{n(E)}{n(S)}.
\]

**Example 4.9.** An urn contains five red balls and four white balls. A sample of two balls is selected at random from the urn.

(a) Find the probability that both balls are red.

(b) Find the probability that at least one of the balls is white.

*(Solution.*) The sample space for this experiment consists of all possible ways to select 2 balls from the 9 balls in the urn. Since the order in which the balls are selected doesn’t matter, there are \( C(9,2) = \frac{9!}{2!7!} = 36 \) possible ways to make this selection, so the sample space has size 36. We’ll use this in both parts of the problem.

(a) Let \( E \) be the event that both balls are red. For this event to happen, we must choose 2 balls from the 5 red balls in the urn, and there are \( C(5,2) = 10 \) ways to do this. We conclude that

\[
P(E) = \frac{n(E)}{n(S)} = \frac{10}{36} \approx 0.2778.
\]

That is, there is a 0.2778 probability that both balls are red.
(b) Let $F$ be the event that at least one ball is white. There are two ways this could happen: we could draw one white ball and one red ball, or we could draw two white balls. There are $C(4, 1)$ ways to select one white ball from the four, and $C(5, 1)$ ways to select one red ball from the five, so the multiplication principle tells us that there are $C(4, 1) \cdot C(5, 1)$ ways to select one red ball and one white ball. Similar to part (a) of this example, there are $C(4, 2)$ ways to select two white balls from the four white balls in the urn. All told,

$$n(F) = C(4, 1) \cdot C(5, 1) + C(4, 2) = 4 \cdot 5 + \frac{4 \cdot 3}{2!} = 20 + 6 = 26.$$

That is, there are 20 ways to select one white ball and one red ball, and there are 6 ways to select two white balls, so there are 26 ways to select at least one white ball. So we have

$$P(F) = \frac{n(E)}{n(S)} = \frac{26}{36} \approx 0.7222,$$

meaning that there is a 72.22% chance that at least one of the two balls will be white.

\[\square\]

4.3.2 The Complement Rule

Once we’ve done part (a) of Example 4.9, there is a much easier strategy for determining the probability in part (b) than the strategy we employed. Notice that the event “at least one ball is white” is the complement of the event “both balls are red.” That is, if at least one ball is white, then it is not the case that both balls are red, and if it is not the case that both balls are red, then at least one ball must be white. Since the events $E$ and $F$ are complementary, we know that $E \cup F = S$ and $E \cap F = \emptyset$. The principle of inclusion and exclusion then becomes

$$P(S) = P(E) + P(F) - P(\emptyset) \Rightarrow 1 = P(E) + P(F) - 0 \Rightarrow 1 = P(E) + P(F).$$

Because we found in part (a) that $P(E) = \frac{10}{36}$, we see that

$$1 = \frac{10}{36} + P(F) \Rightarrow P(F) = 1 - \frac{10}{36} = \frac{26}{36} \approx 0.7222.$$

This is much easier than finding all the possible ways to draw at least one white ball, and this strategy will be frequently useful.

\[\text{The Complement Rule.} \text{ Let } E \text{ be any event. Then} \]

$$1 = P(E) + P(E').$$

It is often the case that an event’s probability is messy to compute directly, but that the event’s complement has an easily computed probability. In these cases, it makes good sense to compute the probability of the complement and apply the complement rule.
Example 4.10. This example is based on a popular party trick. Suppose we have a group of 50 people. Assuming that birthdays are equally distributed over the 366 possible birthdays, what is the probability that some pair of people in this group share a birthday?

(Solution.) This experiment has an enormous sample space. There are 366 possible birthdays for the first person in the group, 366 possible birthdays for the second person in the group, and so on. The multiplication principle tells us that there are

$$366 \cdot 366 \cdots 366 = 366^{50}$$

possible combinations of birthdays, so $$n(S) = 366^{50}$$. Let $$E$$ be the event “some pair of people in the group share a birthday.” It would be very tedious to compute the number of outcomes in $$E$$ — we would have to find all the possible ways for a pair of people in the group to share a birthday. Instead, let’s consider $$E'$$. This is the event where no two people share a birthday — that is, each birthday in the crowd is unique. This is much easier to compute. The first person in the group can have any of the possible 366 birthdays. The second person in the group can have any of the remaining 365 birthdays — our only condition is that their birthday not be a repeat. The third person can then have any of the remaining 364 birthdays, ensuring that this birthday is not shared with either of the first two people in the group. The fourth person can have any of 363 possible birthdays, and so on. All told, there are

$$n(E') = 366 \cdot 365 \cdot 364 \cdots 317$$

possible ways for no two people in the group to share a birthday. Another way to think of this computation is the following: from the 366 possible birthdays, we need to select 50, and the order in which we select them does matter. Then we have

$$n(E') = P(366, 50) = 366 \cdot 365 \cdot 364 \cdots 317.$$

We can now use the complement rule to compute the probability of $$E$$:

$$P(E) = 1 - \frac{n(E')}{n(S)} = 1 - \frac{366 \cdot 365 \cdots 317}{365^{50}} \approx 0.9701.$$ 

That is, there is approximately a 97% chance that some pair of individuals in the group shares a birthday. This is at least mildly surprising, considering the number of birthdays available. In fact, if the group has 23 or more people, it is more likely than not that some pair of individuals in the crowd shares a birthday.

Example 4.11. Refer to Example 3.20 from Chapter 3. If a shortest route from $$A$$ to $$B$$ is selected at random, what is the probability that this route does not pass through $$C$$?

(Solution.) Computing this probability directly is simply an untenable strategy. We would have to compute the number of shortest paths that do not pass through $$C$$, and then divide this by the number of shortest paths between $$A$$ and $$B$$. But how would we go about
computing the shortest paths that miss $C$ if not by subtracting the paths that do pass through $C$ from the total collection of paths? Instead we use the complement rule. If $E$ is the event “the path misses $C$,” then $E'$ is the event “the path passes through $C$.” We computed in Example 3.20 that $n(E') = 60$, and that $n(S) = 126$, so

$$P(E) = 1 - \frac{n(E')}{n(S)} = 1 - \frac{60}{126} = \frac{66}{126} \approx 0.5238.$$ 

That is, approximately 52.4% of the shortest paths from $A$ to $B$ do not pass through $C$. □

### 4.4 Conditional Probability and Independence

In this section we want to begin investigating the interactions between various events. When an event occurs, the probabilities of some other events will change, while the probabilities of other events will not. For instance, if we choose an individual at random and know that this person is under the age of 12, then we know that the probability of the event “the individual has graduated college” is quite low. That is, the event “the individual is younger than 12” occurred, and this changed the probability of the event “the individual is a college graduate.” The fact that the individual is under the age of 12 does not, on the other hand, impact the probability of the event “the individual is female.” In this section we’ll give names to these relationships between events, and discuss how to compute the changes in probability that occur after an event has occurred. Let’s start with an example.

**Example 4.12.** Suppose 150 individuals are surveyed on their political leanings and their education. Of those surveyed, 76 are Democrats, 92 are college graduates, and 34 are neither Democrats nor college graduates. What is the probability that a randomly selected individual from this group is

(a) both a Democrat and a college graduate?

(b) a Democrat?

(c) a Democrat, given the additional fact that this individual is a college graduate?

**(Solution.** To make things easier on ourselves, let’s put our information into a Venn diagram. If $D$ is the event “Democrat” and $C$ is the event “college graduate”, then we have been given that $n(D) = 76$, $n(C) = 92$, and $n(D' \cap C') = 34$. This last piece of information can be rewritten as $n((D \cup C)') = 34$, which tells us that $n(D \cup C) = 150 - 34 = 116$. Using the inclusion-exclusion principle gives

$$116 = 76 + 92 - n(D \cap C),$$

so $n(D \cap C) = 52$. We can fill this information into our Venn diagram, as in Figure 23, and we’re now ready to answer the questions.
Figure 23: Democrats and college graduates.

(a) Of the 150 respondents, 52 are both Democrats and college graduates, so

\[ P(D \cap C) = \frac{n(D \cap C)}{n(S)} = \frac{52}{150} \approx 0.3467. \]

(b) Of the 150 respondents, 76 are Democrats, so

\[ P(D) = \frac{n(D)}{n(S)} = \frac{76}{150} \approx 0.5067. \]

(c) This question is a bit trickier. Because we know that the respondent is a college graduates, our random selection is among the 92 college graduates. That is, there are only 92 possible outcomes to the experiment “randomly select an individual” once we know that this individual must be a college graduate. Among these college graduates, 52 are Democrats, so the probability of selecting a Democrat given that the individual is a college graduate is \( \frac{52}{92} \approx 0.5652. \)

\[ \square \]

Notice that the probability that a randomly selected individual is a Democrat is higher once we know that the individual is a college graduate. We call the latter probability the **conditional probability** of \( D \), given \( C \). That is, we computed the probability that an individual is a Democrat, subject to the condition that this individual be a college graduate. We denote this conditional probability by \( P(D|C) \). We computed \( P(D|C) \) by dividing the number of college graduates who are Democrats by the number of college graduates. This approach (or something like it) holds more generally.

<table>
<thead>
<tr>
<th>Conditional Probability.</th>
<th>Let ( E ) and ( F ) be events. The conditional probability of ( E ), given ( F ), is given by</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[ P(E</td>
</tr>
<tr>
<td></td>
<td>provided that ( P(F) \neq 0. )</td>
</tr>
</tbody>
</table>
Example 4.13. When a pair of dice is rolled, what is the probability that the sum of the dice is 8, given that the sum is not 7?

(Solution.) Let $E$ be the event that the sum is 8, and let $F$ be the event that the sum is not 7. Then $F'$ is the event that the sum is 7, and among the 36 possible ways to roll two dice, there are six ways that the sum will be 7:

$$F' = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}.$$  

So $P(F') = 6/36 = 1/6$, meaning that $P(F) = 1 - 1/6 = 5/6$. There are five ways for the dice to sum to 8:

$$E = \{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\},$$

so $P(E) = 5/36$. To compute the conditional probability $P(E|F)$ we need to know $P(E \cap F)$. But $E \cap F$ is the event “the sum is 8 and the sum is not 7.” This is the same as the event “the sum is 8,” so $E \cap F = E$ and $P(E \cap F) = P(E) = 5/36$. We’re now ready to compute the conditional probability:

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{5/36}{5/6} = \frac{6}{36} \approx 0.1667.$$

So the probability of rolling a sum of 8 is $\frac{5}{36} \approx 0.1389$, but the probability of rolling a sum of 8 given that the sum is not 7 is $\frac{6}{36} \approx 0.1667$.  

We can rewrite equation (24) by multiplying the $P(F)$ across to obtain another formulation of conditional probability:

<table>
<thead>
<tr>
<th>Product Rule for Conditional Probability. Let $E$ and $F$ be events. Then</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(E \cap F) = P(F) \cdot P(E</td>
</tr>
<tr>
<td>provided $P(F) \neq 0.$</td>
</tr>
</tbody>
</table>

Example 4.14. A grade school is testing its students for a particular type of infection. The probability that a randomly selected child has the infection is 0.15, and a randomly selected infected child has an 80% chance of testing positive for the infection. What is the probability that a randomly selected child tests positive for the infection and actually has the infection?

(Solution.) Let $I$ be the event “the student is infected” and let $P$ be the event “the student tests positive.” We’ve been given that $P(I) = 0.15$ and that $P(P|I) = 0.8$, so the product rule for conditional probability tells us that

$$P(P \cap I) = P(I) \cdot P(P|I) = 0.15 \cdot 0.8 = 0.12.$$ 

That is, 12% of all students will test positive for the infection and actually have it.  

88
Example 4.15. Our class has 15 students: 7 female, 8 male. If we select a sequence of two students at random (without replacement), what is the probability that both are male?

(Solution.) Let \( E_1 \) be the event that the first student is male, and let \( E_2 \) be the event that the second student is male. Since 8 of the 15 students are male, we have

\[
P(E_1) = \frac{8}{15},
\]

Now suppose the first student selected is male. Then 14 students remain, and among these 7 are male. So the conditional probability of \( E_2 \) given \( E_1 \) is

\[
P(E_2|E_1) = \frac{7}{14} = \frac{1}{2}.
\]

We can now apply equation (25) to see that

\[
P(E_1 \cap E_2) = P(E_1) \cdot P(E_2|E_1) = \frac{8}{15} \cdot \frac{1}{2} = \frac{4}{15}.
\]

That is, the probability of both students being male is \( \frac{4}{15} \). \( \square \)

We mentioned at the beginning of this section that we often have events \( E_1 \) and \( E_2 \) so that if \( E_1 \) occurs, then the probability of \( E_2 \) occurring is unaffected. This has not been the case so far (“college graduate” impacted “Democrat,” “sum is not 7” impacted “sum is 8,” and so on), but certainly it is sometimes the case. For the canonical example, suppose we are going to roll two dice (one after another). Before beginning, we find the probability of rolling a 4 on the second roll to be \( \frac{1}{6} \). After our first roll (no matter what it is), the probability of rolling a 4 on the second roll is still \( \frac{1}{6} \). The outcome of our first roll has no impact on the outcome of our second roll, so we say that the two experiments are independent.

To be more explicit, let’s use \( E \) to denote the event “the first roll is a 2” and let’s denote the event “the second roll is a 4” by the letter \( F \). Then \( P(F) = \frac{1}{6} \) and, since our first roll doesn’t impact our second roll, \( P(F|E) = \frac{1}{6} \). That is, the probability of \( F \) occurring is unaffected by the fact that \( E \) has occurred. From this we see that

\[
P(E \cap F) = P(E) \cdot P(F|E) = P(E) \cdot P(F),
\]

according to the product rule for conditional probability. This equation will hold whenever \( E \) and \( F \) are independent events, and it serves as the first of our three independence equations.
Independence Equations. Let $E$ and $F$ be events, and assume that both have nonzero probability. Then $E$ and $F$ are independent if

$$P(E \cap F) = P(E) \cdot P(F).$$

If $E$ and $F$ are independent, then

$$P(E|F) = P(E) \quad \text{and} \quad P(F|E) = P(F).$$

Conversely, if $P(E|F) = P(E)$ and $P(F|E) = P(F)$, then $E$ and $F$ are independent.

Example 4.16. We claimed above that the events $E =$ “the first roll is a 2” and $F =$ “the second roll is a 4” are independent. Verify that this is the case.

(Solution.) We immediately have $P(E) = \frac{1}{6}$ and $P(F) = \frac{1}{6}$, so we’ll work on finding $P(E \cap F)$. There are 36 possible outcomes in rolling the dice: 6 outcomes for the first die times 6 outcomes for the second. Only 1 of these 36 is the outcome (2, 4), so

$$P(E \cap F) = \frac{n(E \cap F)}{n(S)} = \frac{1}{36}.$$ 

This means that

$$P(E \cap F) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = P(E) \cdot P(F),$$

so $E$ and $F$ are indeed independent events.

Example 4.17. Let $E$ and $F$ be independent events with $P(E) = 0.5$ and $P(F) = 0.6$. Find $P(E \cup F)$.

(Solution.) Since $E$ and $F$ are independent, we can use our first independence equation to compute

$$P(E \cap F) = P(E) \cdot P(F) = 0.5 \cdot 0.6 = 0.3.$$ 

From here the principle of inclusion and exclusion gives $P(E \cup F) = P(E) + P(F) - P(E \cap F) = 0.5 + 0.6 - 0.3 = 0.8.$

Example 4.18. Suppose a family has three children. Let $E$ be the event “the oldest child is a boy,” and let $F$ be the event “exactly one of the children is a boy.” Are $E$ and $F$ independent?

(Solution.) Because each child can be either a boy or a girl, there are 8 possible outcomes:

$$S = \{BBB, BBG, BGB, BGG, GBB, GBG, GGB, GGG\}.$$ 

We assume that each outcome has equal probability. Now

$$E = \{BBB, BBG, BGB, BGG\} \quad \text{and} \quad F = \{BGB, BGG, GBG\},$$
so $P(E) = \frac{4}{8} = \frac{1}{2}$ and $P(F) = \frac{3}{8}$. Since 
\[ E \cap F = \{BGG\}, \]
$P(E \cap F) = \frac{1}{8}$. Then 
\[ P(E) \cdot P(F) = \frac{1}{2} \cdot \frac{3}{8} = \frac{3}{16} \neq \frac{1}{16} = P(E \cap F), \]
so $E$ and $F$ are not independent events.

It’s possible that a collection of more than two events can be independent, meaning that none of the events affect any of the other events. In this case, we say that the collection of events is \textbf{pairwise independent}, or just \textbf{independent}, and the first independence equation continues to hold.

\begin{center}
\textbf{Independence Equation for Several Events.} A collection of events $E_1, E_2, \ldots, E_n$ is \textbf{(pairwise)} independent if (and only if)
\[ P(E_1 \cap E_2 \cap \cdots \cap E_n) = P(E_1) \cdot P(E_2) \cdot \cdots P(E_n). \]
\end{center}

\textbf{Example 4.19.} Serena has a job interview tomorrow morning, and she is very paranoid about oversleeping. To combat this, she sets one alarm each on her alarm clock, cell phone, and tablet. Each of these alarms has a 99\% chance of sounding, and the alarms are independent. What is the probability that at least one of the three alarms will sound?

\textit{Proof.} Let $E_1$ be the event that the first alarm sounds, let $E_2$ be the event that the second alarm sounds, and let $E_3$ be the event that the third alarm sounds. Then 
\[ P(E_1) = 0.99, \quad P(E_2) = 0.99, \quad \text{and} \quad P(E_3) = 0.99. \]
From this we could use the independence equation to compute the probability that all three alarms sound: $P(E_1 \cap E_2 \cap E_3) = 0.99 \cdot 0.99 \cdot 0.99$, but this is not what we want. We want the probability that at least one of the alarms will sound, which is $P(E_1 \cup E_2 \cup E_3)$. For this it will be helpful to use the complement rule:
\[ P(E_1 \cup E_2 \cup E_3) = 1 - P((E_1 \cup E_2 \cup E_3)') = 1 - P(E_1' \cap E_2' \cap E_3'). \]
Because we know $P(E_1)$, $P(E_2)$, and $P(E_3)$, we can easily compute $P(E_1')$, $P(E_2')$, and $P(E_3')$:
\[ P(E_1') = 0.01, \quad P(E_2') = 0.01, \quad \text{and} \quad P(E_3') = 0.01. \]
The independence of the events then tells us that 
\[ P(E_1' \cap E_2' \cap E_3') = P(E_1') \cdot P(E_2') \cdot P(E_3') = 0.01^3 = 0.000001; \]
this is the probability that \textit{none} of the alarms go off. Finally, we have 
\[ P(E_1 \cup E_2 \cup E_3) = 1 - P(E_1' \cap E_2' \cap E_3') = 1 - 0.000001 = 0.999999, \]
so there is a 99.9999\% chance that at least one of the alarms will sound. \hfill \Box

91
4.5 Tree Diagrams

In this section we won’t introduce any new mathematical ideas, but we will give a tool that can be very useful in computing probabilities: **tree diagrams.** Problems in probability can often be very difficult to visualize. Tree diagrams can help us to better see how the events involved in a problem fit together. We’ll start by reworking an example from the Section 4.4 group work, this time with the aid of a tree diagram.

**Example 4.20.** A basketball player makes each free throw with a probability of 0.6 and is on the line for a one-and-one free throw. (That is, the second shot is attempted only if the first shot is made.) Assume that the two shots are independent. What is the most likely result: 0 points, 1 point, or 2 points?

*(Solution.)* First, let’s diagram what could happen. The player will either make or miss her first shot. If the first shot is missed, the process is finished. If the first shot is made, the player attempts another shot, which she will either make or miss.

![Tree Diagram](image)

We can now add the relevant probabilities to this diagram. The probability that the first shot is made is 0.6, so the probability that the first shot is missed is 0.4. In the second level of the tree, the first shot has been made and we are interested in the probability that the second shot will be made. Because the shots are independent, the second shot will be made with probability 0.6, and the second shot will be missed with probability 0.4. Altogether we have

![Enhanced Tree Diagram](image)

From this tree we see that there are three outcomes: the player will either miss the first shot (resulting in 0 points), make the first shot and miss the second (resulting in 1 point), or make both shots (resulting in 2 points). The probabilities are then given by

\[
P(0 \text{ points}) = 0.4, \quad P(1 \text{ point}) = 0.6 \cdot 0.4 = 0.24, \quad P(2 \text{ points}) = 0.6 \cdot 0.6 = 0.36.
\]

So the likeliest outcome is that the player will score 0 points, followed by 2 points, and then 1 point.  

\[\square\]
Tree diagrams are especially useful when the events involved are not independent (so that the probability of a given event depends on the events that have already happened). The next example demonstrates this.

**Example 4.21.** We have a blue bag which contains 4 blue chips and 5 orange chips, and we have an orange bag which contains 12 blue chips and 13 orange chips. We randomly select one chip from the blue bag and set it aside. We then draw one chip from the bag that matches this chip. What is the probability that we draw one chip of each color?

*(Solution.)* There’s a lot going on in this problem, and we’ll have a hard time visualizing it without a tree diagram. First let’s diagram what happens without worrying yet about probabilities.

If the first chip we draw is blue, then we make another selection from the blue bag. Because we don’t replace the chip we drew, the blue bag now has 3 blue chips and 5 orange chips. On the other hand, if the first chip is orange, we draw our second chip from the orange bag, which has 12 blue chips and 13 orange chips. We can now add the probabilities to our diagram:
There are two ways to draw one chip of each color: we can draw a blue chip and then an orange chip, or we could draw orange then blue. The probability of blue then orange is \(\frac{4}{9} \cdot \frac{5}{8} = \frac{5}{18}\), and the probability of orange then blue is \(\frac{5}{9} \cdot \frac{12}{25} = \frac{4}{15}\), so the probability of drawing one chip of each color is

\[
\frac{5}{18} + \frac{4}{15} = \frac{25}{90} + \frac{24}{90} = \frac{49}{90} \approx 0.5444.
\]

Both of the examples we’ve dealt with so far have the structure “do one task and then do another.” We either took a free throw and then another, or we drew a chip from a bag, and then drew another chip. The next example demonstrates that tree diagrams can be useful in examples which don’t have this feature.

**Example 4.22.** A certain grade school has grades kindergarten through fifth grade. The kindergarten students account for 17.5% of the student population, and three-fifths of the kindergarten students ride the bus to school. Among the first through fifth-graders, two fifths ride the bus. What is the probability that a randomly selected child rides the bus to school?

*(Solution.)* There are two sets of children that ride the bus to school: the kindergarten bus-riders, and the non-kindergarten bus-riders. Let’s say \(B\) is the event “bus-rider” and \(K\) is the event “kindergartener.” Then the bus-riders consist of \(K \cap B\) (the kindergarten bus-riders) and \(K' \cap B\) (the non-kindergarten bus-riders). Let’s draw a tree to see how these sets fit together:

\[
P(K \cap B) = 0.175 \cdot \frac{3}{5} = 0.105
\]

\[
P(K \cap B') = 0.175 \cdot \frac{2}{5} = 0.07
\]

\[
P(K' \cap B) = 0.825 \cdot \frac{2}{5} = 0.33
\]

\[
P(K' \cap B') = 0.825 \cdot \frac{3}{5} = 0.495
\]

So there is a 10.5% chance that a randomly selected student is a bus-riding kindergartener, and a 33% chance that a randomly selected student is a bus-riding non-kindergartener. Altogether, the probability that a child rides the bus is

\[
P(B) = P(K \cap B) + P(K' \cap B) = 0.105 + 0.33 = 0.435,
\]

so 43.5% of the children in the school ride the bus. \(\square\)
4.6 Bayes’ Theorem, Natural Frequencies

Let’s return to an example from the Section 4.5 group work.

**Example 4.23.** At a local college, four sections of economics are taught during the day and five sections are taught at night. Seventy-five percent of the day sections are taught by full-time faculty. Forty percent of the evening sections are taught by full-time faculty. If Janet has a part-time teacher for her economics course, what is the probability that she is taking a night class?

*(Solution.*) Let $D$, $N$, $F$, and $E$ denote the events “day class,” “night class,” “full time professor,” and “part-time professor,” respectively. Then the problem statement provides $P(D)$, $P(N)$, $P(F|D)$, and $P(F|N)$, and asks us to compute $P(N|E)$. First we recall that

$$P(N|E) = \frac{P(N \cap E)}{P(E)}$$

and then use a tree diagram to compute $P(N \cap E)$ and $P(E)$. Our tree diagram is

![Tree Diagram](https://via.placeholder.com/150)

From this we read off that

$$P(N \cap E) = P(N) \cdot P(E|N) = \frac{5}{9} \cdot 0.6 = \frac{3}{5}.$$  

To find $P(E)$ we must add $P(D \cap E)$ to $P(N \cap E)$. We have

$$P(E) = P(N \cap E) + P(D \cap E) = P(N) \cdot P(E|N) + P(D) \cdot P(E|D) = \frac{3}{9} + \frac{1}{9} = \frac{4}{9}.$$  

Finally we compute

$$P(N|E) = \frac{P(N \cap E)}{P(E)} = \frac{P(N) \cdot P(E|N)}{P(N) \cdot P(E|N) + P(D) \cdot P(E|D)} = \frac{3/9}{4/9} = \frac{3}{4}.$$
so the probability that Janet is taking a night class is 0.75.

This example is our first brush with Bayes’ theorem. In computing $P(E)$ we had to find all the paths through our tree that resulted in $E$. It turned out that there were two — $N$ then $E$ and $D$ then $E$ — and the sum of the probabilities of these two events is the probability of $E$. One of these two paths — $N$ then $E$ — gives $N \cap E$, and we divide this probability by $P(E)$ to find $P(N|E)$. This holds whenever $B_1$ and $B_2$ are mutually exclusive events so that $B_1 \cup B_2 = S$ (in which case we say that $B_1$ and $B_2$ partition $S$).

Bayes’ Theorem. If $B_1$ and $B_2$ are mutually exclusive events so that $B_1 \cup B_2 = S$, then for any event $A$,

$$P(B_1|A) = \frac{P(B_1) \cdot P(A|B_1)}{P(B_1) \cdot P(A|B_1) + P(B_2) \cdot P(A|B_2)}.$$

If $B_1, B_2,$ and $B_3$ are mutually exclusive events so that $B_1 \cup B_2 \cup B_3 = S$, then for any event $A$,

$$P(B_1|A) = \frac{P(B_1) \cdot P(A|B_1)}{P(B_1) \cdot P(A|B_1) + P(B_2) \cdot P(A|B_2) + P(B_3) \cdot P(A|B_3)}.$$

More generally, if the events $B_1, B_2, \ldots, B_n$ partition the sample space $S$, then

$$P(B_1|A) = \frac{P(B_1) \cdot P(A|B_1)}{P(B_1) \cdot P(A|B_1) + P(B_2) \cdot P(A|B_2) + \cdots + P(B_n) \cdot P(A|B_n)}$$

for any event $A$.

The only good way to get a handle on the formulas in Bayes’ theorem is to work examples, so that’s what we’ll do now.

Example 4.24. The following table gives the distribution of income and shows the proportion of two-car families by income level for a certain suburban county:

<table>
<thead>
<tr>
<th>Annual Income</th>
<th>Proportion of pop.</th>
<th>% with 2 or more cars</th>
</tr>
</thead>
<tbody>
<tr>
<td>$&lt;30,000$</td>
<td>0.10</td>
<td>20</td>
</tr>
<tr>
<td>$30,000 - 44,999$</td>
<td>0.20</td>
<td>50</td>
</tr>
<tr>
<td>$45,000 - 59,999$</td>
<td>0.35</td>
<td>60</td>
</tr>
<tr>
<td>$60,000 - 74,999$</td>
<td>0.30</td>
<td>75</td>
</tr>
<tr>
<td>$\geq75,000$</td>
<td>0.05</td>
<td>90</td>
</tr>
</tbody>
</table>

Suppose that a randomly chosen family has two or more cars. What is the probability that its income is at least $75,000? (Solution.) Let’s label the five income brackets with the following event names: $B_1, B_2, B_3, B_4,$ and $B_5$. That is, we are going to select a family at random, and $B_3$ is the event “this family’s annual income is between $45,000 and $59,999.” Then the table gives the following
probabilities:

\[ P(B_1) = 0.1, \quad P(B_2) = 0.2, \quad P(B_3) = 0.35, \quad P(B_4) = 0.3, \quad P(B_5) = 0.05. \]

Now we let \( A \) be the event “the family owns 2 or more cars.” The table also gives the following conditional probabilities:

\[ P(A|B_1) = 0.2, \quad P(A|B_2) = 0.5, \quad P(A|B_3) = 0.6, \quad P(A|B_4) = 0.75, \quad P(A|B_5) = 0.9. \]

We can now apply Bayes’ theorem to find \( P(B_5|A) \):

\[
P(B_5|A) = \frac{P(B_5) \cdot P(A|B_5)}{P(B_1) \cdot P(A|B_1) + \cdots + P(B_5) \cdot P(A|B_5)}
= \frac{0.1 \cdot 0.2 + 0.2 \cdot 0.5 + 0.35 \cdot 0.6 + 0.3 \cdot 0.75 + 0.05 \cdot 0.9}{0.045}
= \frac{0.02 + 0.1 + 0.21 + 0.225 + 0.045}{0.6} \approx 0.075.
\]

So given that a family owns two or more cars, there is a 7.5% chance that this family has an annual income of at least $75,000.

Example 4.25. Your two roommates, Alex and Pat, each take a morning shower. Alex sings in the shower 50% of the time, and Pat sings in the shower 90% of the time. One morning you hear the water running and whoever is in the shower is singing. What is the probability that Pat is taking a shower?

(Solution.) Let \( B_1 \) be the event “Pat is taking a shower,” and let \( B_2 \) be the event “Alex is taking a shower.” The event “someone is singing in the shower” will be denoted \( A \). We assume that, a priori, \( B_1 \) and \( B_2 \) are equally likely, so that \( P(B_1) = P(B_2) = 0.5 \). Because Pat sings in the shower 90% of the time, \( P(A|B_1) = 0.9 \); since Alex sings in the shower 50% of the time, \( P(A|B_2) = 0.5 \). We can now use Bayes’ theorem to compute the probability that Pat is in the shower, given that we hear singing:

\[
P(B_1|A) = \frac{0.5 \cdot 0.9}{0.5 \cdot 0.9 + 0.5 \cdot 0.5} = \frac{0.45}{0.7} \approx 0.6429.
\]

So there is a 64.3% chance that the voice we hear belongs to Pat.
5 Probability and Statistics

In this chapter we draw on discussions and examples found in [1].

5.1 Visual Representations of Data

In this section we’ll discuss a few different ways that data can be presented. First we want to discuss bar charts and histograms, both of which we can construct from the even more basic frequency table. For instance, consider Table 1, which gives the 2014 populations of the four American states that share a border with Mexico. This is the frequency table for

<table>
<thead>
<tr>
<th>State</th>
<th>Population</th>
<th>% of total population</th>
</tr>
</thead>
<tbody>
<tr>
<td>California</td>
<td>38.8 million</td>
<td>52.0</td>
</tr>
<tr>
<td>Arizona</td>
<td>6.73 million</td>
<td>9.0</td>
</tr>
<tr>
<td>New Mexico</td>
<td>2.09 million</td>
<td>2.8</td>
</tr>
<tr>
<td>Texas</td>
<td>26.96 million</td>
<td>36.2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>74.58 million</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

Table 1: States bordering Mexico.

the given data, because it shows the frequency with which each response occurs. From this table we see, for instance, that 52% of people who live in an American state that shares a border with Mexico live in California, while only 2.8% of such people live in New Mexico. Another natural way to present this data is as a bar chart, such as the one seen in Figure 24a. From this bar chart we can see that the population of California is vastly greater than that of Arizona or New Mexico, and significantly greater even than the population of Texas. Adjacent to Figure 24a we have Figure 24b, another bar chart representing the same data. But in this chart the y-axis is not labeled with the populations of the various states, but with each state’s proportion of the total population of Mexico-bordering states. Since the shapes of the charts are the same, we get the same feel for the data from either presentation; they differ only in which statistic about each state is being represented.

<table>
<thead>
<tr>
<th>State</th>
<th>% of total population</th>
<th>Sector angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>California</td>
<td>52.0</td>
<td>187.2°</td>
</tr>
<tr>
<td>Arizona</td>
<td>9.0</td>
<td>32.4°</td>
</tr>
<tr>
<td>New Mexico</td>
<td>2.8</td>
<td>10.08°</td>
</tr>
<tr>
<td>Texas</td>
<td>36.2</td>
<td>130.32°</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>100.0</td>
<td><strong>360°</strong></td>
</tr>
</tbody>
</table>

Table 2: Sector angles for states bordering Mexico.
From the bar charts given in 24 we can see that the populations of California and Texas account for the great majority of the people living in Mexico-bordering states. Another way to depict this proportion is with a pie chart, such as the one in Figure 25. We come up with such a chart by allotting to each of the four states a sector of our circle (or pie) which is commensurate to the proportion of this states’ populations. The last column of Table 2 gives the angle for each sector. This column is computed by multiplying each relevant percentage by the 360° in our circle. This gives the angle of the sector of our pie which will be associated to the given state. Because California accounts for over half of the population, the sector associated to California has an angle of over 180°.

We created the bar charts in Figure 24 by noticing that each data point (that is, person who lives in a Mexico-bordering state) falls into one of four bins (that is, states). We plotted the number of data points in each bin and called the resulting chart a bar chart. When our bins are numerical rather than qualitative things like states, we call our bar chart a
**Histogram.** For example, consider the following homework scores:

$$4, 10, 4, 8, 9, 10, 0, 6, 9, 8, 7, 0, 7, 9, 10.$$ 

From this data we can construct a frequency table:

<table>
<thead>
<tr>
<th>Score</th>
<th># of occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
</tr>
</tbody>
</table>

and we can also provide a bar chart, which is known in this case as a histogram. This histogram for this data is seen in Figure 26.

There’s one more way to represent data that we’d like to mention, called the **box plot**. We use box plots when we have a set of measurements or other numbers and we want to see how these are distributed from least to greatest. To create a box plot, we first compute what is called the **five-number summary** for our data. The five numbers that make up the five-number summary are

- the **minimum** of the data set.
- the **first quartile** of the data set. This is a number below which 25% of the data points lie.
- the **median** or **second quartile** of the data set. Half of the data points are less than the median, and half are greater than it. (Approximately. There may be some points which are equal to the median.)
- the **third quartile** of the data set. Similar to the first quartile, but 75% of the data points lie below this number, with 25% above.
- the **maximum** of the data set.

To demonstrate, let’s find the five-number summary for the homework scores given above. It’s best if we first arrange them in increasing order:

$$0, 0, 4, 4, 6, 7, 7, 8, 8, 9, 9, 9, 10, 10, 10.$$
This list has 15 data points, so the eighth point in the list is the median. The median homework score is 8: 7 students earned a score less than 8 and 7 students earned a score greater than 8. One way to think of the first quartile is that it is the median of the numbers less than the median. There are 7 numbers less than the median, and they are

\[0, 0, 4, 4, 6, 7, 7.\]

The median of this list is its fourth data point, so the first quartile is 4. Finally, the third quartile is the median of those numbers above the median, which are

\[8, 9, 9, 9, 10, 10, 10.\]

Notice that we have an eight on this list, even though 8 is the median. This is because the score 8 appears twice in our first list. The first 8 was our median, and this 8 lies to the right of the median for the purposes of our computation. The median of this smaller collection is 9, so our third quartile is 9. Altogether,

\[\text{min} = 0, \quad Q_1 = 4, \quad Q_2 = 8, \quad Q_3 = 9, \quad \text{max} = 10.\]

We can now use this five-number summary to produce a box plot, such as the one seen in Figure 27.

We produce such a box plot by drawing “fences” at each of the numbers in our five-number summary, and then closing up the \(Q_1\), \(Q_2\), and \(Q_3\) fences into a box. Half of our data points lie between \(Q_1\) and \(Q_3\), so half of our data points are inside this box. Within the box, half of the data points lie on either side of the median, so the line inside the box splits the data so that half points lie on either side. We often call the width of this box the \textbf{interquartile range}. In this case, the interquartile range is \(Q_3 - Q_1 = 9 - 4 = 5.\)

### 5.2 Frequency and Probability Distributions

Suppose we’re trying to decide whether Chris Paul or Jimmy Butler is the better scorer in the NBA, based on Paul’s scoring in the 2015-2016 season and Butler’s scoring in the
<table>
<thead>
<tr>
<th>Points scored</th>
<th>Number of Occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chris Paul</td>
</tr>
<tr>
<td>0-9</td>
<td>3</td>
</tr>
<tr>
<td>10-19</td>
<td>36</td>
</tr>
<tr>
<td>20-29</td>
<td>30</td>
</tr>
<tr>
<td>30-39</td>
<td>4</td>
</tr>
<tr>
<td>40-49</td>
<td>1</td>
</tr>
<tr>
<td>50+</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>74</td>
</tr>
</tbody>
</table>

Table 3: Scoring frequencies. Source: Basketball-Reference.

2014-2015 and 2015-2016 seasons. Some of this data is presented in Table 3. Each row represents a scoring bracket, and the corresponding entry in each player’s column tells us how many times the player scored in that bracket. For instance, Chris Paul scored 20-29 points 30 times in the 2015-2016 season, while Jimmy Butler scored in this bracket 56 times over the course of the 2014-2015 and 2015-2016 seasons. The table given here is called a frequency distribution, because it tells us the frequency with which each outcome (in our case, scoring bracket) occurred.

Our task now is to compare the scoring tendencies of Chris Paul and Jimmy Butler. One difficulty with this is the fact that the frequency data is taken over different time periods. It’s not really fair to say that Jimmy Butler scores 20-29 points almost twice as often as Chris Paul, because the data was taken over a longer period. Instead, we compute the relative frequency distributions for the two players. To find the relative frequency with which an outcome occurred, we divide the total number of instances of this outcome by the total number of outcomes. In Chris Paul’s case, he scored between 20 and 29 points in 30 different games, over the course of 74 games. This means that he scored 20-29 points in 40.54% of his games. On the other hand, Jimmy Butler scored in this bracket a total of 56 times over the course of 130 games, good for 43.07% of his games. So while Jimmy Butler scored in the 20-29 bracket almost twice as many times as Chris Paul, his advantage isn’t as dramatic when we compare relative frequencies. The relative frequency data for both players is tabulated in Table 4.

Once the data sets have been normalized in this way, we can create visual representations which will admit meaningful comparisons. For instance, we can now use the data in Table 4 to create histograms for the scoring tendencies of Chris Paul and Jimmy Butler, and it will make sense to compare these histograms.

In Figure 28 we have histograms for Chris Paul (on the left) and Jimmy Butler (on the right) comparing their frequencies for the various scoring brackets. From this data we can see that Jimmy Butler seems to score at a slightly higher rate than Chris Paul — he has
the taller bar in the higher scoring ranges — but his advantage is not nearly as great as it looked when we saw the frequency distribution.

The height of Chris Paul’s 10-19 points bar is 0.4865, and its width is 1. So the area of this bar is 0.4865, which corresponds precisely with the fact that 48.65% of Chris Paul’s games see him score 10-19 points. We can add to this the area of his 0-9 points bar (an area of 0.0405) to see that Chris Paul scores fewer than 20 points in 48.65% + 4.05% = 52.7% of his games. In fact this holds more generally: the frequency with which an event happens corresponds to the area this event takes up in a frequency histogram.

Example 5.1. The following is the frequency histogram for Jimmy Butler:
In what proportion of his games does Jimmy Butler score between 10 and 29 points?

(Solution.) The games in which Jimmy Butler scores 10-29 points are represented by the blue bars in this chart. The areas of these bars are 0.3846 and 0.4308, respectively, so their total area is 0.8154. So Jimmy Butler scores between 10 and 29 points in 81.54% of his games.

This area summation game isn’t only useful for histograms of frequency data. We also like to use it for histograms of probability distributions. For example, consider the experiment of rolling two dice and observing their sum. We have computed before that this experiment has the following probability distribution:

<table>
<thead>
<tr>
<th>Sum</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>(\frac{1}{36})</td>
<td>(\frac{2}{36})</td>
<td>(\frac{3}{36})</td>
<td>(\frac{4}{36})</td>
<td>(\frac{5}{36})</td>
<td>(\frac{6}{36})</td>
<td>(\frac{5}{36})</td>
<td>(\frac{4}{36})</td>
<td>(\frac{3}{36})</td>
<td>(\frac{2}{36})</td>
<td>(\frac{1}{36})</td>
</tr>
</tbody>
</table>

From this probability distribution we can generate the histogram seen in Figure 29a. As is the case for frequency data, areas correspond to probabilities. For instance, the area of the blue region in Figure 29b is

\[
\frac{5}{36} + \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{15}{36},
\]

and this is the probability that the sum of the two dice will be greater than 7.

We’ll often denote the outcome of an experiment by the letter \(X\), or by another capital letter. We say that the outcome is a random variable, an we discuss the probability that our random variable takes different values. For instance, in our two-dice experiment, \(X\) would be the sum of the two dice and our probability distribution would be

<table>
<thead>
<tr>
<th>(k)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P(X = k))</td>
<td>(\frac{1}{36})</td>
<td>(\frac{2}{36})</td>
<td>(\frac{3}{36})</td>
<td>(\frac{4}{36})</td>
<td>(\frac{5}{36})</td>
<td>(\frac{6}{36})</td>
<td>(\frac{5}{36})</td>
<td>(\frac{4}{36})</td>
<td>(\frac{3}{36})</td>
<td>(\frac{2}{36})</td>
<td>(\frac{1}{36})</td>
</tr>
</tbody>
</table>

The primary advantage of this notation is that it keeps us from having to repeat things like “the two dice add up to 8.” We can just say “\(X = 8\)” Another advantage is that random variables can be manipulated according to the algebraic rules we already know. For example, suppose the random variable \(X\) has the following probability distribution:
Then we could define a new random variable $Y$ by setting $Y = X^2$. This random variable can only take on three different values: 0, 1, or 4, depending on the value of $X$. If $X = \pm 2$, then $Y = 4$, and this happens with probability $0.2 + 0.1 = 0.3$. Similarly, $Y = 1$ whenever $X = \pm 1$, so $P(Y = 1) = 0.4 + 0.2 = 0.6$. Finally, $Y = 0$ exactly when $X = 0$, so $P(Y = 0) = 0.1$. So the probability distribution for $Y$ is given by

<table>
<thead>
<tr>
<th>$k$</th>
<th>$P(Y = k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
<td>0.6</td>
</tr>
<tr>
<td>2</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Let’s finish this section with one more example.

**Example 5.2.** Consider the following probability distribution:

<table>
<thead>
<tr>
<th>$k$</th>
<th>$P(Z = k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3/15</td>
</tr>
<tr>
<td>1</td>
<td>2/15</td>
</tr>
<tr>
<td>2</td>
<td>4/15</td>
</tr>
<tr>
<td>3</td>
<td>5/15</td>
</tr>
<tr>
<td>4</td>
<td>1/15</td>
</tr>
</tbody>
</table>

(The probability that $Y = 4$ is intentionally left blank.)

(a) Determine the probability that $Z = 4$.

(b) Find $P(Z \geq 2)$.

(c) Find the probability that $Z$ is at most 3.

(d) Find the probability that $Z + 2$ is less than 4.

(e) Find the probability distribution of $Y = (Z - 2)^2$.

(Solution.)

(a) From the table we see that the possible outcomes for $Z$ are 0, 1, 2, 3, and 4. The sum of the probabilities of these outcomes must be 1, so we conclude that

\[ P(Z = 4) = 1 - P(Z = 0) - P(Z = 1) - P(Z = 2) - P(Z = 3) = 1 - \frac{3}{15} - \frac{2}{15} - \frac{4}{15} - \frac{5}{15} = \frac{1}{15}. \]
(b) There are a couple of ways we can do this. The event $Z \geq 2$ consists of three outcomes: $Z = 2$, $Z = 3$, and $Z = 4$. We can sum the probabilities of these outcomes to conclude that

$$P(Z \geq 2) = P(Z = 2) + P(Z = 3) + P(Z = 4) = \frac{4}{15} + \frac{5}{15} + \frac{1}{15} = \frac{10}{15}.$$  

Alternatively, we can recognize that the complement of the event $Z \geq 2$ is $Z < 2$, an event with two outcomes: $Z = 0$ and $Z = 1$. Then

$$P(Z \geq 2) = 1 - P(Z < 2) = 1 - \frac{3}{15} - \frac{2}{15} = \frac{10}{15}.$$  

In this case the choice of whether or not to use the complement rule makes little difference, but we can often save ourselves a lot of time by recognizing situations where the complement consists of fewer outcomes.

(c) The phrase “$Z$ is at most 3” can be written as $Z \leq 3$. The complement of this event is $Z > 3$, an event with just one outcome: $Z = 4$. So

$$P(Z \leq 3) = 1 - P(Z > 3) = 1 - P(Z = 4) = \frac{14}{15}.$$  

So the probability that $Z$ is at most 3 is 14/15.

(d) This is a good opportunity to practice the algebraic manipulations that random variable notation allows. We write “$Z + 2$ is less than 4” as $Z + 2 < 4$, and then rewrite this as $Z < 2$. This is an event with two outcomes: $Z = 0$ and $Z = 1$. So

$$P(Z + 2 < 4) = P(Z = 0) + P(Z = 1) = \frac{3}{15} + \frac{2}{15} = \frac{5}{15}.$$  

So the probability is 1/3 that $Z + 2$ is less than 4.

(e) Again the random variable notation comes in handy. Since the possible outcomes for $Z$ are 0, 1, 2, 3, and 4, the possible outcomes for $Y$ are 0, 1, and 4. Notice that $Y = 0$ can be written as

$$0 = (Z - 2)^2 \quad \Rightarrow \quad 0 = Z - 2 \quad \Rightarrow \quad Z = 2.$$  

So $P(Y = 0) = P(Z = 2) = 4/15$. Similarly we have $Y = 1$ when

$$1 = (Z - 2)^2 \quad \Rightarrow \quad \pm 1 = Z - 2 \quad \Rightarrow \quad Z = 2 \pm 1.$$  

So $P(Y = 1) = P(Z = 1) + P(Z = 3) = 2/15 + 5/15 = 7/15$. Finally, if $Y = 4$, then

$$4 = (Z - 2)^2 \quad \Rightarrow \quad \pm 2 = Z - 2 \quad \Rightarrow \quad Z = 2 \pm 2.$$  

So $P(Y = 4) = P(Z = 0) + P(Z = 4) = 3/15 + 1/15 = 4/15$. We summarize our work by giving the probability distribution of $Y$:

<table>
<thead>
<tr>
<th>$k$</th>
<th>0</th>
<th>1</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(Y = k)$</td>
<td>$\frac{4}{15}$</td>
<td>$\frac{7}{15}$</td>
<td>$\frac{4}{15}$</td>
</tr>
</tbody>
</table>

□
5.3 Binomial Trials

A probability problem of frequent real-world interest is the **binomial trial**. The general setup of a binomial trial is that we have some task we want to attempt or test we want to apply, and each time we attempt this task or apply this test the probability of success is \( p \). For example, \( p \) might be:

- the probability that a baseball player gets a hit in a given at-bat;
- the probability that a device is found to be defective upon testing;
- the probability that a mortgage will be defaulted upon.

Notice that the word “success” is not meant to assign any value of goodness to an outcome — defective devices and defaulted mortgages are bad — but is simply the word we use for the outcome we are trying to measure. What we’re really interested in is the following question: if we attempt this task or apply this test \( n \) times (and the attempts/tests are independent), what is the probability of exactly \( k \) successes? That is, we care about questions like:

- if a baseball player has four at-bats, what is the probability he will have two hits?
- if we test 15 items, what is the probability that none will be defective?
- if we administer 100 mortgages, what is the probability that fewer than ten will be defaulted upon?

These questions all have the same structure, so we would like to find a general procedure that will handle them. We start with an example.

**Example 5.3.** The Charleston Music Group produces guitar strings, and these strings must meet certain design specifications. The strings meet these specifications with probability 0.96 (that is, 96% of all strings produced meet the specifications). If an inspector selects 10 strings to test, what is the probability that all 10 strings meet the specifications? What is the probability that exactly 9 of the strings meet the specifications?

*(Solution.)* The first event — that all 10 strings pass the test — is actually the intersection of 10 independent events: the event that string 1 passes the test, the event that string 2 passes the test, and so on. Each of these events has probability \( p = 0.96 \), and the strings are assumed to be independent, so the probability that all 10 events occur is

\[
(0.96)(0.96)(0.96)(0.96)(0.96)(0.96)(0.96)(0.96)(0.96)(0.96) \approx 0.6648.
\]

Next we have the event that exactly nine strings pass the test. From the 10 strings we’re testing, we must choose nine which will pass the test, and there are \( C(10, 9) = 10 \) ways to do this. Each of these 10 outcomes has the same probability: each one is the intersection of 9 successes (in some order) and 1 failure, and thus has probability

\[
p^9 q^1.
\]
where \( p = 0.96 \) is the probability of success and \( q \) is the probability of failure. Since failure is the complement of success\(^ {14} \), \( q = 1 - p = 0.04 \). So there are 10 outcomes in which exactly nine strings fail the test, and each of these 10 outcomes has probability

\[
(0.96)^9 \cdot (0.04)^1 \approx 0.0277.
\]

The probability of exactly nine successes is then

\[
10 \cdot 0.0277 = 0.277,
\]

so approximately 27.7\% of the time, exactly nine strings will pass the test. \(\Box\)

Our work in Example 5.3 leads us to some important observations about binomial trials. The most obvious is that in any binomial trial, \( q = 1 - p \), since \( q \) is the probability of failure, and failure is the complement of success. Next we see that if we’re counting successes in \( n \) independent trials, there are a number of ways for, say, \( k \) successes to occur. From the \( n \) trials, we must choose \( k \) which are successes, so there are \( C(n, k) \) outcomes with exactly \( k \) successes. Finally, Example 5.3 pointed out that each of the \(15\) \( C(n, k) \) outcomes with \( k \) successes have the same probability. These outcomes have \( k \) successes and \( n - k \) failures, and thus have probability

\[
p^k q^{n-k}.
\]

Adding up all of these outcomes, the probability of exactly \( k \) successes is \( \binom{n}{k} p^k q^{n-k} \). Let’s summarize our observations about binomial trials:

**Binomial Trials.** Suppose we perform \( n \) independent, binomial trials, and that the random variable \( X \) counts the number of successes\(^ {16} \) in these trials. (We call such a random variable a **binomial random variable**.) If each trial succeeds with probability \( p \) (and thus fails with probability \( q = 1 - p \)), then

\[
P(X = k) = \binom{n}{k} p^k q^{n-k},
\]

for \( k = 0, 1, 2, \ldots, n \).

**Example 5.4.** We roll a single die 7 times. We consider each roll to be a success if either a 3 or a 5 appears, and to be a failure otherwise. What is the probability of exactly 6 successes? What is the probability of at least one success?

\(^{14}\) “Failure is the complement of success.” - Wayne Gretzky - Michael Scott

\(^{15}\) Remember that \( C(n, k) \) and \( \binom{n}{k} \) mean the same thing.

\(^{16}\) Calling our outcomes “successes” and “failures” can lead to some confusion. Sometimes \( X \) is counting the number of **bad** outcomes (say, testing positive for an infection), but we still say that \( p \) is the probability of success.
(Solution.) Let $X$ be the number of successes. Since 2 of the 6 possible outcomes of a single roll are considered successes, $p = 2/6 = 1/3$, and thus $q = 2/3$. The probability of exactly 6 successes is given by

$$P(X = 6) = \binom{7}{6} \left(\frac{1}{3}\right)^6 \left(\frac{2}{3}\right)^{7-6} = \frac{7!}{6!1!} \cdot \frac{12}{2187} = \frac{14}{2187} \approx 0.0064.$$ 

Next we want to compute the probability of at least one success. One way to find this probability is to sum the probabilities of all the outcomes that make up this event:

$$P(X \geq 1) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6) + P(X = 7).$$

A better way is to use the complement rule. The complement of the event “at least one success” is the event “no successes”. Thus

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{7}{0} \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^7 = 1 - \frac{128}{2187} \approx 1 - 0.0585 = 0.9415.$$

So there is a 94.15% chance that at least one of our rolls will be either a 3 or a 5. □

Example 5.5. A hitting coach makes a friendly bet with his best home run hitter: if the player hits a home run in today’s game, the coach will buy the player a steak. Otherwise, the player must buy the coach a steak. Suppose the player has a 10% chance of hitting a home run in any given at-bat, and that the player takes 5 at-bats in today’s game. What is the probability that the coach will have to buy dinner? Give the probability distribution for the number of home runs the player will hit.

(Solution.) Let $X$ be the number of home runs that the player hits in today’s game. As long as $X$ is not 0, the coach must buy the player’s dinner. Since the player has 5 at-bats in today’s game, $X$ is allowed to take any of six different values: 0, 1, 2, 3, 4, or 5. The probability that $X = 0$ is given by

$$P(X = 0) = \binom{5}{0} (0.1)^0 (0.9)^5 = (0.9)^5 \approx 0.59049,$$

where $p = 0.1$ and $q = 0.9$, because the probability of a home run in any given at bat is 0.1. At this point we can find the probability that the coach will have to buy dinner. The coach only buys if $X \geq 1$, so we can use the complement rule to find that

$$P(X \geq 1) = 1 - P(X = 0) \approx 1 - 0.59049 = 0.40951.$$ 

So the coach must buy dinner with probability 0.40951. To give the full probability distribution for $X$ we compute the probabilities of the remaining outcomes:

$$P(X = 1) = \binom{5}{1} (0.1)^1 (0.9)^4 = 5(0.1)(0.9)^4 \approx 0.32805$$

109
\[ P(X = 2) = \binom{5}{2} (0.1)^2 (0.9)^3 = 10(0.1)^2 (0.9)^3 \approx 0.0729 \]
\[ P(X = 3) = \binom{5}{3} (0.1)^3 (0.9)^2 = 10(0.1)^3 (0.9)^2 \approx 0.0081 \]
\[ P(X = 4) = \binom{5}{4} (0.1)^4 (0.9)^1 = 5(0.1)^4 (0.9) \approx 0.00045 \]
\[ P(X = 5) = \binom{5}{5} (0.1)^5 (0.9)^0 = (0.1)^5 \approx 0.00001 \]

The probability distribution for \( X \) is then given by

<table>
<thead>
<tr>
<th>( k )</th>
<th>( P(X = k) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.59049</td>
</tr>
<tr>
<td>1</td>
<td>0.32805</td>
</tr>
<tr>
<td>2</td>
<td>0.0729</td>
</tr>
<tr>
<td>3</td>
<td>0.0081</td>
</tr>
<tr>
<td>4</td>
<td>0.00045</td>
</tr>
<tr>
<td>5</td>
<td>0.00001</td>
</tr>
</tbody>
</table>

With this probability distribution in hand, we now have a second way to compute the probability that the coach buys dinner:

\[ P(X \geq 1) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) \approx 0.40951. \]

Just for fun, here’s a histogram of this probability distribution:

Notice that the probabilities for 4 home runs and 5 home runs do not result in a discernible bar. As it happens, only 17 players have hit 4 home runs in a single game in the history of Major League Baseball, and no player has ever hit 5 home runs in a single game. \( \square \)

### 5.4 The Mean

In almost all cases where statistics are used, we are hoping to estimate various properties of a population by properties of a sample. A population is some set about which information
is desired, and a sample is some randomly selected subset of the population\textsuperscript{17}. For example, if we want to study the political preferences of women over the age of 40, our population will be all women over the age of 40 and we might survey 1,200 women over 40 on their political preferences. These 1,200 women comprise our sample, and we will use the properties of this sample (party preference, voter turnout, etc.) to estimate the properties of the population.

Perhaps the most basic sample statistic (and certainly one of the most widely-reported) is the \textbf{sample mean}. If we measure some quantity about all the elements of our sample, the average of all these measurements is called the sample mean, and is denoted $\bar{x}$.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|}
\hline
\textbf{Sample Mean.} The \textbf{sample mean} of a sample of $n$ measurements $x_1, x_2, \ldots, x_n$ is \\
\hline
$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$ \\
\hline
\end{tabular}
\end{table}

\textbf{Example 5.6.} Suppose we want to estimate the average age of all Pepperdine students by asking 20 students for their age. The ages of these 20 students are

\begin{center}
18, 20, 21, 19, 22, 19, 20, 21, 19, 20, 19, 22, 19, 21, 18, 27, 20, 21, 18.
\end{center}

What is the sample mean for this sample?

\textit{(Solution.)} We have

\begin{align*}
\bar{x} &= \frac{1}{20} (18 + 20 + 21 + 19 + 22 + 19 + 20 + 21 + 19 + 19 \\
&\quad + 20 + 19 + 22 + 19 + 21 + 18 + 27 + 20 + 21 + 18) \\
&= \frac{403}{20} = 20.15,
\end{align*}

so the sample mean is 20.15 years old. Notice that the actual average age of a Pepperdine student may not be 20.15 years old, but this is the average age of our sample, and we use it to approximate the average age of a Pepperdine student. \hfill \square

Computing sample means is a straightforward task, but we can add (at least) one wrinkle. The data in Example 5.6 is not presented in a particularly nice manner, nor is the resulting computation all that pretty. One way to spruce up the presentation of the data (and thus to complicate the computation) is to present the data in the form of a frequency distribution. For the ages given, the frequency distribution is

\textsuperscript{17}It's important that our samples be randomly selected. Otherwise the properties of the sample may not accurately reflect the properties of the population.
Now when we compute the average age, we know that we will have 3 copies of 18, 6 copies of 19, etc. So our average age is given by

\[
\bar{x} = \frac{1}{20} (3 \cdot 18 + 6 \cdot 19 + 4 \cdot 20 + 4 \cdot 21 + 2 \cdot 22 + 1 \cdot 27) = \frac{403}{20} = 20.15.
\]

So computing the sample mean from the frequency data is perhaps slightly less straightforward than computing it directly from the data, but this computation certainly has a cleaner appearance. Let’s generalize this frequency-based approach.

### Sample Mean from Frequency Data.

Suppose we have a sample of \( n \) elements, and this sample includes \( k \) distinct outcomes: \( x_1, x_2, \ldots, x_k \). Moreover, suppose that the outcome \( x_1 \) occurs \( f_1 \) times (that is, \( f_1 \) is the frequency of \( x_1 \)), that the outcome \( x_2 \) occurs \( f_2 \) times, and so on, so that

\[ f_1 + f_2 + \cdots + f_k = n. \]

Then the sample mean is given by

\[
\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \cdots + x_k f_k}{n}.
\]

Equivalently,

\[
\bar{x} = x_1 \cdot \frac{f_1}{n} + x_2 \cdot \frac{f_2}{n} + \cdots + x_k \cdot \frac{f_k}{n}.
\]

The numbers \( \frac{f_1}{n}, \ldots, \frac{f_k}{n} \) are the relative frequencies of \( x_1, \ldots, x_k \).

### Example 5.7.

We’ve already computed the average age of the students in our sample using the traditional average and using the frequency information. Now compute the sample mean using the relative frequency data.

*(Solution.)* First we find the relative frequency distribution. We do this by dividing each of the frequencies in our frequency distribution by the total number of observations (20):
With this distribution in hand, we compute

\[ \bar{x} = 18 \cdot 0.15 + 19 \cdot 0.30 + 20 \cdot 0.20 + 21 \cdot 0.20 + 22 \cdot 0.10 + 27 \cdot 0.05 = 20.15, \]

just as before. \[\Box\]

Because the sample mean is based on data gathered from a sample rather than the entire population, it is called a **statistic**. The sample mean hopes to approximate the **population mean**, which is found by computing the average age of *all* Pepperdine students, instead of just the 20 that we sampled. The population mean is denoted by \( \mu \), and is computed in the same way that the sample mean is:

\[ \mu = \frac{x_1 + x_2 + \cdots + x_N}{N}, \]

where in our case \( x_1, x_2, \ldots, x_N \) are the ages of *all* Pepperdine students. We say that \( \mu \) is a **parameter** of the population. We will always use statistics (measurements taken from samples) to approximate parameters (measurements of the population).

When the results of an experiment follow some probability distribution, we can use probability to estimate the mean of these results in advance. If \( X \) is a random variable corresponding to the output of some experiment, then we can compute the **expected value** of \( X \), and this gives a theoretical estimate for the mean of the results of our experiment.

### Expected Value.

Suppose a random variable \( X \) has \( k \) possible outcomes: \( x_1, x_2, \ldots, x_k \). Moreover, suppose that \( P(X = x_1) = p_1 \), that \( P(X = x_2) = p_2 \), and so on. Then the **expected value** of \( X \) is given by

\[ E(X) = x_1p_1 + x_2p_2 + \cdots + x_kp_k. \]

The expected value of \( X \) is what we expect the mean to be if an experiment is done a large number of times; in particular, the expected value of \( X \) is frequently not actually an outcome that \( X \) could assume, as the following example demonstrates.

**Example 5.8.** Consider the experiment of rolling a fair die, and let \( X \) be the number showing on the die. Compute \( E(X) \).
The probability distribution for $X$ is

<table>
<thead>
<tr>
<th>$k$</th>
<th>$P(X = k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1/6$</td>
</tr>
<tr>
<td>2</td>
<td>$1/6$</td>
</tr>
<tr>
<td>3</td>
<td>$1/6$</td>
</tr>
<tr>
<td>4</td>
<td>$1/6$</td>
</tr>
<tr>
<td>5</td>
<td>$1/6$</td>
</tr>
<tr>
<td>6</td>
<td>$1/6$</td>
</tr>
</tbody>
</table>

From this we compute

$$E(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{21}{6} = 3.5,$$

so the expected value of $X$ is 3.5.

Note that we can’t actually roll a 3.5, but this is what we expect the average roll to be after a large number of rolls. For instance, consider the following 30 rolls, randomly generated by WolframAlpha:

6, 3, 6, 4, 1, 4, 4, 6, 2, 2, 3, 3, 5, 6, 5, 2, 3, 1, 6, 2, 4, 2, 1, 3, 4, 6, 4, 6, 4.

We immediately notice that none of these values is the expected value of 3.5. However, the average value of these 30 rolls is $56/15 \approx 3.7333$, which is quite close to our expected value of 3.5. After generating another 30 rolls, say

1, 3, 5, 3, 5, 3, 6, 3, 5, 1, 2, 2, 6, 2, 6, 2, 5, 3, 5, 3, 2, 5, 5, 4, 5, 5, 2, 4, 1, 4,

the running average is $11/3 \approx 3.6667$, still closer to the expected value of 3.5.

**Example 5.9.** Suppose we flip a coin 4 times and let $Y$ count the number of tails that appear. How many tails should we expect?

*(Solution.*) Our intuition immediately leads us to an answer here: since heads and tails are equally likely, we expect about half of our flips to be tails, so our expected value should be $4/2 = 2$. This turns out to be the correct answer, but for the purposes of illustration, let’s compute this the hard way. First we’ll compute the probability distribution for $Y$. We know that $Y$ is a binomial random variable with parameters $n = 4$ and $p = 0.5$ (meaning that we have 4 trials, with probability of success 0.5), so

<table>
<thead>
<tr>
<th>$k$</th>
<th>$P(Y = k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\binom{4}{0}(0.5)^0(0.5)^4 = 1/16$</td>
</tr>
<tr>
<td>1</td>
<td>$\binom{4}{1}(0.5)^1(0.5)^3 = 1/4$</td>
</tr>
<tr>
<td>2</td>
<td>$\binom{4}{2}(0.5)^2(0.5)^2 = 3/8$</td>
</tr>
<tr>
<td>3</td>
<td>$\binom{4}{3}(0.5)^3(0.5)^1 = 1/4$</td>
</tr>
<tr>
<td>4</td>
<td>$\binom{4}{4}(0.5)^4(0.5)^0 = 1/16$</td>
</tr>
</tbody>
</table>

114
From here we can compute
\[ E(Y) = 0 \cdot \frac{1}{16} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{16} = \frac{1}{4} + \frac{3}{4} + \frac{3}{4} + \frac{1}{4} = 2. \]
So the expected value is 2 tails, as we predicted.

Binomial random variables come up a lot, so here’s a formula to cut straight to their expected values:

**Expected value of a binomial random variable.** Suppose \( X \) is a binomial random variable with parameters \( n \) and \( p \). Then \( E(X) = np \).

**Example 5.10.** A very accurate college basketball player makes 88% of her free throws. If she attempts 175 free throws over the course of a season, how many should she expect to make (assuming free throws are independent)?

*(Solution.)* Let \( X \) be a random variable that counts the successful free throws. Then \( X \) is a binomial random variable with parameters \( n = 175 \) and \( p = 0.88 \), so
\[ E(X) = (175)(0.88) = 154. \]
That is, we expect the player to make 154 of her free throws over the course of the season.

We’ll conclude this section with a relatively important applications of expected values.

**Example 5.11.** Using life insurance tables, a retired man determines that the probability that he will live 5 more years is 0.9. He decides to take out a life insurance policy that will pay $10,000 in the event that he dies during the next 5 years. How much should he be willing to pay for this policy, disregarding such factors as interest rates or inflation?

*(Solution.)* Let \( x \) be the amount of money that the man pays for the life insurance policy, measured in dollars, and let the random variable \( X \) be the amount of money paid to the man or his family by the life insurance policy, also measured in dollars. If the man does not die in the next 5 years, \( X = -x \), since he has paid the life insurance company \( x \) dollars. If the man does die, \( X = 10,000 - x \), because the life insurance policy pays $10,000 to the man’s family, and we subtract the amount paid by the man in the first place. So \( X \) has the following probability distribution:

<table>
<thead>
<tr>
<th>Amount</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-x)</td>
<td>0.9</td>
</tr>
<tr>
<td>$(10,000 - x)\</td>
<td>0.1</td>
</tr>
</tbody>
</table>

From this we compute
\[ E(X) = (-x)(0.9) + (10,000 - x)(0.1) = 1,000 - x. \]
The man does not want the expected value of this transaction to be negative\(^{18}\), so he should not pay more than $1,000 for the life insurance policy.

\(^{18}\)For insurance customers, we will see that variance is a more important measure than expected value.
5.5 The Variance and Standard Deviation

In Section 5.4 we introduced one of the primary goals of statistics: approximating parameters. Recall that the mean is a parameter of a population; from a population we can take a sample and compute the mean of this sample. This sample mean is a statistic, and as our sample size grows, this statistic should converge to the parameter we’re trying to measure — the mean. We’re now ready to introduce two new parameters of interest: the variance and the standard deviation. Both of these parameters intend to measure the extent to which our outcomes bunch up near or spread out around the mean.

For instance, consider the experiments “roll a die” and “roll two dice and observe their sum.” If we conduct these experiments enough times and create histograms of our outcomes, we expect the histograms to look like those seen in Figure 30. In both cases, the expected value is given by a vertical blue line, and we see that the outcomes are evenly distributed about these lines. But in the case of rolling a single die, the outcomes are spread out — we are just as likely to roll a 6, which is 2.5 units away from the mean, as we are to roll a 4, which is 0.5 units away from the mean. In the case of rolling two dice and taking their sum, events that are near the expected value are more likely than the more extreme values.

In real life we’re often willing to trade a situation with high variance for one that has lower variance, despite having a less desirable expected value. For instance, in Example 5.11 we considered a man buying a $10,000, 5-year life insurance policy, and this man suspects that he has a probability of 0.9 of living at least 5 more years. If he pays $1,200 for the life insurance policy, then his expected outcome is

$$E(X) = (0.9)(-1,200) + (0.1)(8,800) = -200.$$  

That is, if he repeats this process over and over, he expects that he will, on average, lose $190. But the next five years of his life are an experiment this man will only conduct once. Perhaps he’s willing to accept the likelihood that he will lose money in exchange for the guarantee that his family won’t be stuck with $10,000 of expenses after his death. Without the insurance, the man’s expected outcome is

$$E(X) = (0.9)(0) + (0.1)(-10,000) = -100,$$

Figure 30: Probability distributions with their means.
assuming he expects the costs associated with his death to total $10,000. Not purchasing insurance gives the man a better expected value — he expects, on average, for his family to pay $100 instead of $2000. But perhaps the lower variance associated with buying the insurance is worth the lower expected value.

Let’s start trying to define variance. We want to measure the extent to which our outcomes “spread out” and take values far away from the mean. A first attempt at measuring deviation from the mean might be to simply look at the difference between our outcomes and the mean. That is, if we have measurements \( x_1, \ldots, x_N \), with mean \( \mu \), perhaps we want to look at the average of the differences between these measurements and the mean:

\[
\frac{(x_1 - \mu) + (x_2 - \mu) + \cdots + (x_N - \mu)}{N}.
\]

The problem with this approach is that it doesn’t distinguish between negative and positive differences. Extreme values at the low end will cancel out extreme values at the high end. Instead, we square the differences between our measurements and the mean:

<table>
<thead>
<tr>
<th>Population Variance and Standard Deviation.</th>
<th>Suppose we have a population ( x_1, x_2, \ldots, x_N ) whose mean is ( \mu ). Then the variance of this population is</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma^2 = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \cdots + (x_N - \mu)^2}{N} )</td>
<td>Alternatively we could define ( \sigma^2 ) in terms of the relative frequencies of ( x_1, \ldots, x_N ). If the outcome ( x_1 ) occurs with frequency ( f_1 ), ( x_2 ) with frequency ( f_2 ), and so on up to ( f_r ), then</td>
</tr>
<tr>
<td>( \sigma^2 = (x_1 - \mu)^2 \cdot \frac{f_1}{N} + (x_2 - \mu)^2 \cdot \frac{f_2}{N} + \cdots + (x_r - \mu)^2 \cdot \frac{f_r}{N} )</td>
<td>In either case, the standard deviation of the population is given by ( \sigma = \sqrt{\sigma^2} ).</td>
</tr>
</tbody>
</table>

Notice that the standard deviation doesn’t measure anything that wasn’t already measured by the variance. The reason for caring about the standard deviation is that it has the same units as our measurements, where the variance has these units squared.

**Example 5.12.** A women’s basketball team has 11 players, and their heights are as follows (in inches):

\[58, 59, 57, 59, 57, 56, 58, 59, 58, 59, 58.\]

Find the variance of the heights of the players.

*(Solution.)* First we compute the population mean:

\[
\mu = \frac{58 + 59 + 57 + 59 + 57 + 56 + 58 + 59 + 58 + 59 + 58}{11} = 58.
\]
Then we have

\[ \sigma^2 = \frac{1}{11} \left[ (58 - 58)^2 + (59 - 58)^2 + (57 - 58)^2 + (59 - 58)^2 + (57 - 58)^2 \\
+ (56 - 58)^2 + (58 - 58)^2 + (59 - 58)^2 + (58 - 58)^2 + (59 - 58)^2 + (58 - 58)^2 \right] \]

\[ = \frac{1}{11} \left[ 0 + 1 + 1 + 1 + 4 + 0 + 1 + 0 + 1 + 0 \right] = \frac{10}{11} \approx 0.9091. \]

This means that the standard deviation of this population is \( \sigma = \sqrt{\frac{10}{11}} \approx 0.9535 \) inches. □

In Section 5.4 we computed the sample mean and the population mean, and the formulas for these two were identical. The sample mean is a statistic, while the population mean is a parameter, and we say that the sample mean is unbiased, because as we take larger and larger samples the sample mean will tend towards the population mean. Unfortunately, in order to ensure that our sample variance (that is, the variance of a sample taken from a population) is an unbiased estimator for the population variance, we have to tweak our formula a bit.

**Sample Variance and Standard Deviation.** Suppose we have a sample \( x_1, x_2, \ldots, x_n \) taken from a population, and the mean of this sample is \( \bar{x} \). Then the variance of this sample is

\[ s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n-1}. \]

Alternatively, if the outcome \( x_1 \) occurs with frequency \( f_1 \), \( x_2 \) with frequency \( f_2 \), and so on up to \( f_r \), then

\[ s^2 = (x_1 - \bar{x})^2 \cdot \frac{f_1}{n-1} + (x_2 - \bar{x})^2 \cdot \frac{f_2}{n-1} + \cdots + (x_r - \bar{x})^2 \cdot \frac{f_r}{n-1}. \]

In either case, the standard deviation of the sample is given by \( s = \sqrt{s^2} \).

**Example 5.13.** Consider the sample of Pepperdine students taken in Example 5.6. Compute the variance of this sample of ages.

*(Solution.)* The sample in Example 5.6 included 20 students, and we tabulated the frequency data for the six outcomes in that example. We also computed the average age \( \bar{x} = 20.15 \) of this sample. Based on the frequency data and mean observed there, we have

\[ s^2 = (18 - 20.15)^2 \cdot \frac{3}{19} + (19 - 20.15)^2 \cdot \frac{6}{19} + (20 - 20.15)^2 \cdot \frac{4}{19} \]

\[ + (21 - 20.15)^2 \cdot \frac{4}{19} + (22 - 20.15)^2 \cdot \frac{2}{19} + (27 - 20.15)^2 \cdot \frac{1}{19} \approx 4.1342. \]
This means that \( s = \sqrt{s^2} \approx 2.0333 \), so most students’ ages should be within 2 or so years of the average age of 20.15 years. □

As we did with the mean, we would also like to define the variance for probability distributions. This will allow us to predict the variance or standard deviation of a collection of outcomes by computing these parameters on the probability distribution that we expect the outcomes to follow. The formula for variance will not come as a surprise.

**Variance.** The **variance** of a random variable \( X \) is defined by

\[
V(X) = E[(X - \mu)^2],
\]

where \( \mu = E(X) \) is the expected value of \( X \). As with the population and sample variances, the **standard deviation** of a random variable \( X \), denoted \( \sigma_X \), is given by

\[
\sigma_X = \sqrt{V(X)}.
\]

**Example 5.14.** Suppose a random variable \( X \) has the following probability distribution:

<table>
<thead>
<tr>
<th>( k )</th>
<th>(-20)</th>
<th>(-10)</th>
<th>(0)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P(X = k) )</td>
<td>0.1</td>
<td>0.2</td>
<td>0.4</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Compute \( V(X) \).

*(Solution.)* First we compute the mean \( \mu \):

\[
\mu = (-20)(0.1) + (-10)(0.2) + (0)(0.4) + (10)(0.3) = -1.
\]

Now we have the following table:

<table>
<thead>
<tr>
<th>( k )</th>
<th>( P(X = k) )</th>
<th>( k - \mu )</th>
<th>((k - \mu)^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-20</td>
<td>0.1</td>
<td>-19</td>
<td>361</td>
</tr>
<tr>
<td>-10</td>
<td>0.2</td>
<td>-9</td>
<td>81</td>
</tr>
<tr>
<td>0</td>
<td>0.4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>0.3</td>
<td>11</td>
<td>121</td>
</tr>
</tbody>
</table>

We can now compute the expected value of \((X - \mu)^2\):

\[
V(X) = 361 \cdot 0.1 + 81 \cdot 0.2 + 1 \cdot 0.4 + 121 \cdot 0.3 = 89.
\]

So the variance of \( X \) is 89, and the standard deviation is \( \sigma_X = \sqrt{89} \approx 9.4340 \). □

Variance computations can get out of hand pretty quickly. Now we’ll give a second way to compute variance that is sometimes less messy. We won’t verify that this expression is equivalent.
Alternate Expression for Variance. For any random variable $X$,

$$V(X) = E(X^2) - [E(X)]^2.$$  

**Example 5.15.** Compute the variance for the random variable in Example 5.14 using the alternate expression.

*(Solution.)* We’ve already computed that $E(X) = -1$. Next we compute

$$E(X^2) = (400)(0.1) + (100)(0.2) + (0)(0.4) + (100)(0.3) = 40 + 20 + 30 = 90.$$

Then

$$V(X) = E(X^2) - [E(X)]^2 = 90 - (-1)^2 = 89,$$

exactly as we found above. □

**Example 5.16.** Consider the experiment of rolling a die, and let $X$ be the number observed. Compute $V(X)$.

*(Solution.)* The probability distribution for $X$ is given by

<table>
<thead>
<tr>
<th>$k$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>P($X = k$)</td>
<td>1/6</td>
<td>1/6</td>
<td>1/6</td>
<td>1/6</td>
<td>1/6</td>
<td>1/6</td>
</tr>
</tbody>
</table>

We have previously computed $E(X) = 3.5$, and we can compute from this distribution that

$$E(X^2) = 1 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 9 \cdot \frac{1}{6} + 16 \cdot \frac{1}{6} + 25 \cdot \frac{1}{6} + 36 \cdot \frac{1}{6} = \frac{91}{6},$$

so

$$V(X) = \frac{91}{6} - \left( \frac{21}{6} \right)^2 = \frac{35}{12} \approx 2.9167.$$

So $X$ has a variance of about 2.9167, and thus has a standard deviation of $\sigma_X \approx 1.7078$. □

It’s important to note that the units of the standard deviation match the units of $X$. As a number, the variance of a probability distribution seems relatively meaningless. We know that a greater variance indicates that the outcomes are more spread out, but what can we make of the fact that the variance in the experiment of rolling a single die is 2.92? The standard deviation is helpful because its units actually match those of the random variable. When we say that the standard deviation of $X$ is about 1.71 (as we do in the die-rolling example), we mean that most outcomes are within 1.71 units of the expected outcome of 3.5. Indeed, the only outcomes not within one standard deviation of 3.5 are 1 and 6.

When we introduced the expected value of a random variable in Section 5.4, we made special mention of the expected value of a binomial random variable. A binomial random
variable with parameters \( n \) and \( p \) has expected value \( np \). Now we point out the variance (and thus also the standard deviation) of a binomial random variable.

<table>
<thead>
<tr>
<th>Variance of a Binomial Random Variable.</th>
<th>Suppose ( X ) is a binomial random variable with parameters ( n ) and ( p ). Then</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V(X) = npq ),</td>
<td></td>
</tr>
<tr>
<td>and thus</td>
<td></td>
</tr>
<tr>
<td>( \sigma_X = \sqrt{npq} ),</td>
<td></td>
</tr>
<tr>
<td>where ( q = 1 - p ).</td>
<td></td>
</tr>
</tbody>
</table>

**Example 5.17.** Suppose we flip a coin five times and let \( Y \) count the number of tails. Which outcomes are within one standard deviation of the expected value?

*(Solution.)* Our parameters are \( n = 5 \) and \( p = 0.5 \), so first we compute the expected value:

\[
E(Y) = 5(0.5) = 2.5.
\]

So we expect 2.5 tails in the five flips. Our variance is

\[
V(Y) = 5(0.5)(0.5) = 1.25,
\]

so \( \sigma_Y = \sqrt{1.25} \approx 1.1180 \). The outcomes of 2 tails and 3 tails are both within one standard deviation of the expected value, but the outcomes of 0, 1, 4, or 5 tails are all more than 1.1180 tails away from 2.5, and thus outside of our desired range. \( \square \)

We’ve remarked on a few occasions in this section that we expect most of our outcomes to lie within a single standard deviation of the mean. Indeed, given some lower and upper bounds on measurements, we’d often like to approximate the number of outcomes that fall within this range. In the case of probability distributions, we can compute a (rough) lower bound for the probability that an outcome lies within a certain distance of the mean. Unfortunately we cannot justify it here.

<table>
<thead>
<tr>
<th>Chebychev’s Inequality.</th>
<th>Consider a probability distribution with expected value ( \mu ) and standard deviation ( \sigma ). The probability that a randomly chosen outcome lies within ( \mu - c ) and ( \mu + c ) is at least ( 1 - \left( \frac{\sigma}{c} \right)^2 ). Symbolically,</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P(\mu - c \leq X \leq \mu + c) \geq 1 - \left( \frac{\sigma}{c} \right)^2 ),</td>
<td></td>
</tr>
<tr>
<td>where ( X ) is a random variable following the probability distribution.</td>
<td></td>
</tr>
</tbody>
</table>

**Example 5.18.** Consider the experiment of flipping a coin 100 times. Give an upper bound for the probability that we flip fewer than 10 heads or more than 90 heads.
(Solution.) Let $X$ count the number of heads. Then $\mu = np = 100 \cdot 0.5 = 50$ and $\sigma = \sqrt{npq} = \sqrt{100 \cdot 0.5 \cdot 0.5} = 5$. We want the probability that $X < 10$ or $X > 90$. The complement of this event is the event $10 \leq X \leq 90$, and Chebychev’s inequality tells us that

$$P(10 \leq X \leq 90) \geq 1 - \left( \frac{5}{40} \right)^2 = 1 - \frac{1}{64} = \frac{63}{64}.$$ 

Since the event $10 \leq X \leq 90$ has probability at least $63/64$, the event $X < 10$ or $X > 90$ has probability at most $1/64 \approx 0.0156$. □
References

