

# Applied Finite Mathematics

Sekhon, Bloom, Manlove

April 17, 2024



# Contents

<b>1</b>	<b>Linear Functions</b>	<b>1</b>
1.1	Graphing a Linear Equation . . . . .	1
1.1.1	Graphing a Line from its Equation . . . . .	1
1.1.2	Intercepts: . . . . .	4
1.1.3	Graphing a Line from Its Equation in Parametric Form . . . . .	5
1.1.4	Horizontal and Vertical Lines . . . . .	6
1.2	Slope of a Line . . . . .	7
1.3	Determining the Equation of a Line . . . . .	13
1.4	Applications . . . . .	17
1.5	More Applications . . . . .	21
1.5.1	Finding the Point of Intersection of Two Lines . . . . .	21
1.5.2	Solving Systems of Equations: The Elimination Method . . . . .	22
1.5.3	Supply, Demand, and the Equilibrium Market Price . . . . .	24
1.5.4	Break-Even Point . . . . .	25
<b>2</b>	<b>Matrices</b>	<b>29</b>
2.1	Introduction to Matrices . . . . .	29
2.1.1	Vocabulary . . . . .	31
2.1.2	Matrix Addition and Subtraction . . . . .	32
2.1.3	Multiplying a Matrix by a Scalar . . . . .	33
2.1.4	Multiplication of Two Matrices . . . . .	34
2.1.5	Systems of Linear Equations . . . . .	38
2.2	Systems of Linear Equations; Gauss-Jordan Method . . . . .	40
2.2.1	Row Operations in Gauss-Jordan Method . . . . .	43
2.3	Systems of Linear Equations – Special Cases . . . . .	47
2.4	Inverse Matrices . . . . .	51
2.5	Application of Matrices in Cryptography . . . . .	58

2.5.1	Using Matrices for Encoding and Decoding . . . . .	59
2.6	Applications – Leontief Models . . . . .	65
2.6.1	The Closed Model . . . . .	65
2.6.2	The Open Model . . . . .	68
<b>3</b>	<b>Linear Programming with Geometry</b>	<b>73</b>
3.1	Maximization Applications . . . . .	73
3.2	Minimization Applications . . . . .	82
<b>4</b>	<b>Linear Programming, Simplex Method</b>	<b>91</b>
4.1	Linear Programming Applications in Business, Finance, Medicine, and Social Science . . . . .	91
4.1.1	Airline Scheduling . . . . .	93
4.1.2	Kidney Donation Chain . . . . .	94
4.1.3	Advertisements in Online Marketing . . . . .	95
4.1.4	Loans . . . . .	96
4.1.5	Production Planning and Scheduling in Manufacturing . . . . .	96
4.1.6	Bike Share Programs . . . . .	97
4.2	Maximization by the Simplex Method . . . . .	98
4.3	Minimization by the Simplex Method . . . . .	108
<b>5</b>	<b>Exponential and Logarithmic Functions</b>	<b>115</b>
5.1	Exponential Growth and Decay Models . . . . .	115
5.1.1	Comparing Exponential and Linear Growth . . . . .	116
5.1.2	Using Exponential Functions to Model Growth and De- cay . . . . .	117
5.1.3	Comparing Linear, Exponential, and Power Functions . . . . .	122
5.1.4	The Natural Base . . . . .	124
5.2	Graphing Exponential Functions . . . . .	126
5.2.1	Properties of Exponential Growth Functions . . . . .	127
5.2.2	Properties of Exponential Decay Functions . . . . .	128
5.2.3	An Exponential Function Is a One-to-One Function . . . . .	128
5.3	Logarithms and Logarithmic Functions . . . . .	131
5.3.1	Define the Logarithm . . . . .	131
5.3.2	Natural and Base 10 Logarithms . . . . .	133
5.3.3	Properties of Logarithms . . . . .	136
5.4	Graphs and Properties of Logarithmic Functions . . . . .	136

5.5	Application Problems with Exponential and Logarithmic Functions . . . . .	140
5.5.1	Strategies for Solving Equations That Contain Exponents	140
5.5.2	Expressing Exponential Functions in Different Forms . . . . .	147
<b>6</b>	<b>Mathematics of Finance</b>	<b>151</b>
6.1	Simple Interest and Discount . . . . .	151
6.1.1	Simple Interest . . . . .	151
6.1.2	Discounts and Proceeds . . . . .	154
6.2	Compound Interest . . . . .	155
6.2.1	Compound Interest . . . . .	155
6.2.2	Compounding Periods . . . . .	157
6.2.3	Effective Interest Rate . . . . .	160
6.2.4	Continuous Compounding . . . . .	161
6.3	Annuities and Sinking Funds . . . . .	163
6.3.1	Ordinary Annuity . . . . .	163
6.3.2	Sinking Funds . . . . .	167
6.3.3	Annuities Due . . . . .	167
6.4	Present Value of an Annuity and Installment Payment . . . . .	169
6.4.1	Installment Payment on a Loan . . . . .	173
6.4.2	Alternate Formula to find Present Value of an Annuity	174
6.5	Miscellaneous Application Problems . . . . .	175
6.5.1	Outstanding Balance on a Loan . . . . .	176
6.5.2	Problems Involving Multiple Stages of Savings and/or Annuities . . . . .	178
6.5.3	Fair Market Value of a Bond . . . . .	180
6.5.4	Amortization Schedule for a Loan . . . . .	183
6.6	Classification of Finance Problems . . . . .	185
6.6.1	Classification of Problems and Equations for Solutions	186
6.7	Financial Calculations Vocabulary . . . . .	188
<b>7</b>	<b>Sets and Counting</b>	<b>191</b>
7.1	Sets . . . . .	191
7.1.1	Introduction to Sets . . . . .	191
7.1.2	Empty Set, Set Equality, Subsets . . . . .	192
7.1.3	Union of Two Sets . . . . .	192
7.1.4	Intersection of Two Sets . . . . .	192
7.1.5	Complement of a Set, Disjoint Sets . . . . .	192

7.1.6	Venn Diagrams . . . . .	194
7.2	Tree Diagrams and the Multiplication Axiom . . . . .	199
7.2.1	Permutations . . . . .	204
7.3	Circular Permutations and Permutations with Similar Elements	208
7.3.1	Circular Permutations . . . . .	208
7.3.2	Permutations with Similar Elements . . . . .	210
7.4	Combinations . . . . .	212
7.5	Combinations Involving Several Sets . . . . .	217
7.5.1	A Standard Deck of 52 Playing Cards . . . . .	220
7.6	Binomial Theorem . . . . .	221
<b>8</b>	<b>Probability</b>	<b>225</b>
8.1	Sample Spaces and Probability . . . . .	225
8.1.1	Sample Spaces . . . . .	226
8.2	Mutually Exclusive Events and the Addition Rule . . . . .	232
8.3	Probability Using Tree Diagrams and Combinations . . . . .	238
8.3.1	Using Tree Diagrams to Calculate Probabilities . . . . .	238
8.3.2	Using Combinations to Find Probabilities . . . . .	242
8.4	Conditional Probability . . . . .	245
8.5	Independent Events . . . . .	253
<b>9</b>	<b>More Probability and Games of Chance</b>	<b>261</b>
9.1	Binomial Probability . . . . .	261
9.2	Bayes' Formula . . . . .	267
9.3	Expected Value . . . . .	272
9.4	Probability Using Tree Diagrams . . . . .	276
<b>10</b>	<b>Markov Chains</b>	<b>281</b>
10.1	Introduction to Markov Chains . . . . .	281
10.2	Applications of Markov Chains . . . . .	291
10.3	Regular Markov Chains . . . . .	293
10.3.1	Does every Markov chain reach the state of equilibrium?	294
10.3.2	Does the product of an equilibrium vector and its transition matrix always equal the equilibrium vector? . . .	296
10.3.3	Can the equilibrium vector $E$ be found without raising the transition matrix $T$ to large powers? . . . . .	297
10.3.4	Does the long term market share for a Markov chain depend on the initial market share? . . . . .	299

10.4 Absorbing Markov Chains . . . . .	301
10.4.1 Drunkard's Random Walk . . . . .	304
10.4.2 Gambler's Ruin Problem . . . . .	306
<b>11 Game Theory</b>	<b>313</b>
11.1 Strictly Determined Games . . . . .	313
11.2 Non-Strictly Determined Games . . . . .	316
11.3 Reduction by Dominance . . . . .	321





# Chapter 1

## Linear Functions

In this chapter, you will learn to:

1. Graph a linear equation.
2. Find the slope of a line.
3. Determine an equation of a line.
4. Solve linear systems.
5. Do application problems using linear equations.

### 1.1 Graphing a Linear Equation

In this section, you will learn to:

1. Graph a line when you know its equation.
2. Graph a line when you are given its equation in parametric form.
3. Graph and find equations of vertical and horizontal lines.

#### 1.1.1 Graphing a Line from its Equation

Equations whose graphs are straight lines are called linear equations. The following are some examples of linear equations:

$$\begin{aligned}
2x - 3y &= 6, \\
3x &= 4y - 7, \\
y &= 2x - 5, \\
2y &= 3, \\
x - 2 &= 0.
\end{aligned}$$

A line is completely determined by two points. Therefore, to graph a linear equation, we need to find the coordinates of two points. This can be accomplished by choosing an arbitrary value for  $x$  or  $y$  and then solving for the other variable.

**Example 1.1.1.** *Graph the line  $y = 3x + 2$ .*

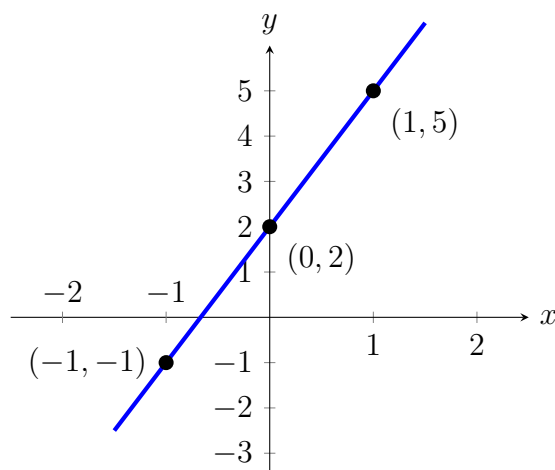
**Solution 1.1.1.** *We need to find the coordinates of at least two points. We arbitrarily choose  $x = -1$ ,  $x = 0$ , and  $x = 1$ .*

*If  $x = -1$ , then  $y = 3(-1) + 2$  or  $y = -1$ . Therefore,  $(-1, -1)$  is a point on this line.*

*If  $x = 0$ , then  $y = 3(0) + 2$  or  $y = 2$ . Hence the point  $(0, 2)$ .*

*If  $x = 1$ , then  $y = 5$ , and we get the point  $(1, 5)$ . Below, the results are summarized, and the line is graphed.*

$x$	$y$
-1	-1
0	2
1	5



**Example 1.1.2.** Graph the line:  $2x + y = 4$

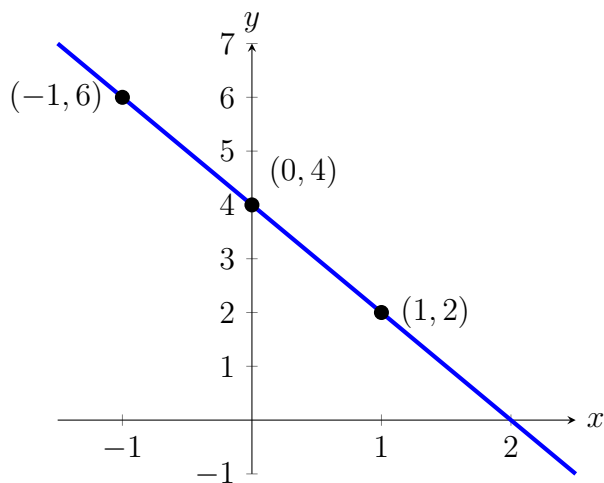
**Solution 1.1.2.** Again, we need to find coordinates of at least two points. We arbitrarily choose  $x = -1$ ,  $x = 0$ , and  $y = 2$ .

If  $x = -1$ , then  $2(-1) + y = 4$  which results in  $y = 6$ . Therefore,  $(-1, 6)$  is a point on this line.

If  $x = 0$ , then  $2(0) + y = 4$ , which results in  $y = 4$ . Hence the point  $(0, 4)$ .

If  $y = 2$ , then  $2x + 2 = 4$ , which yields  $x = 1$ , and gives the point  $(1, 2)$ . The table below shows the points, and the line is graphed.

$x$	$y$
-1	6
0	4
1	2



### 1.1.2 Intercepts:

The points at which a line crosses the coordinate axes are called the intercepts. When graphing a line by plotting two points, using the intercepts is often preferred because they are easy to find.

- To find the value of the x-intercept, we let  $y = 0$ .
- To find the value of the y-intercept, we let  $x = 0$ .

**Example 1.1.3.** Find the intercepts of the line:  $2x - 3y = 6$ , and graph.

**Solution 1.1.3.** To find the x-intercept, let  $y = 0$  in the equation, and solve for  $x$ .

$$2x - 3(0) = 6$$

$$2x = 6$$

$$x = 3$$

Therefore, the x-intercept is the point  $(3, 0)$ .

To find the y-intercept, let  $x = 0$  in the equation, and solve for  $y$ .

$$2(0) - 3y = 6$$

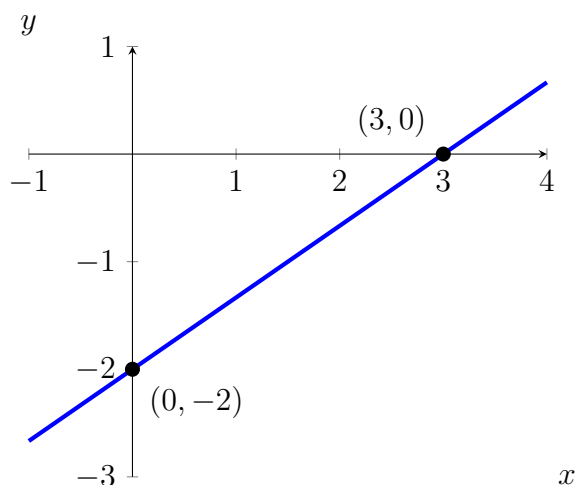
$$0 - 3y = 6$$

$$-3y = 6$$

$$y = -2$$

Therefore, the  $y$ -intercept is the point  $(0, -2)$ .

To graph the line, plot the points for the  $x$ -intercept  $(3, 0)$  and the  $y$ -intercept  $(0, -2)$ , and use them to draw the line.



### 1.1.3 Graphing a Line from Its Equation in Parametric Form

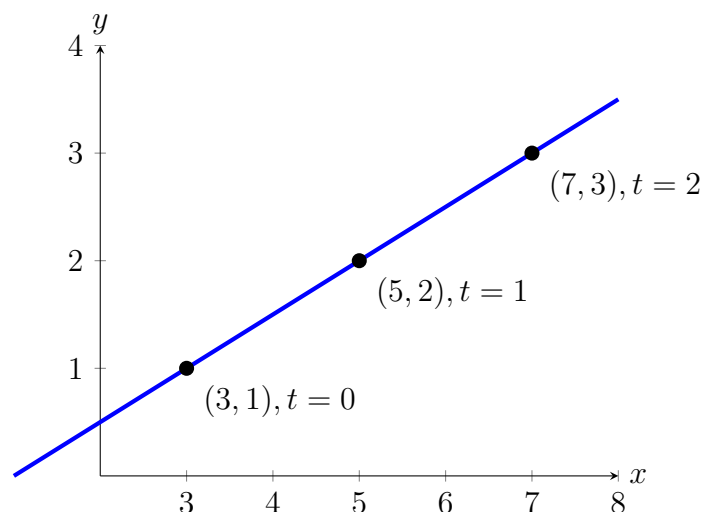
In higher math, equations of lines are sometimes written in parametric form. For example,  $x = 3 + 2t$ ,  $y = 1 + t$ . The letter  $t$  is called the parameter or the dummy variable.

Parametric lines can be graphed by finding values for  $x$  and  $y$  by substituting numerical values for  $t$ . Plot the points using their  $(x, y)$  coordinates and use the points to draw the line.

**Example 1.1.4.** Graph the line given by the parametric equations:  $x = 3 + 2t$ ,  $y = 1 + t$

**Solution 1.1.4.** Let  $t = 0, 1$  and  $2$ ; for each value of  $t$ , find the corresponding values for  $x$  and  $y$ . The results are given in the table below.

$t$	$x$	$y$
0	3	1
1	5	2
2	7	3



### 1.1.4 Horizontal and Vertical Lines

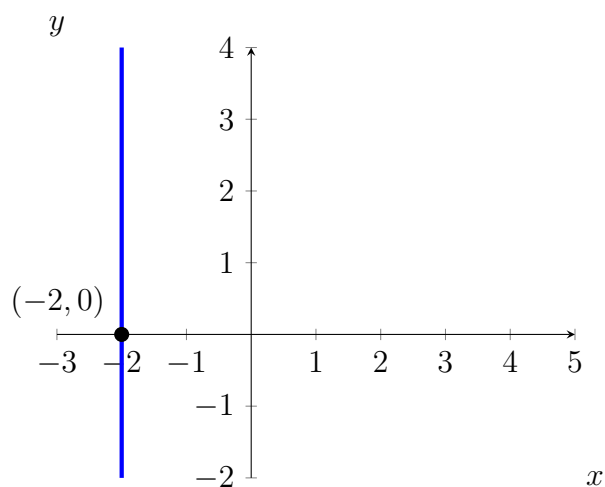
When an equation of a line has only one variable, the resulting graph is a horizontal or a vertical line.

The graph of the line  $x = a$ , where  $a$  is a constant, is a vertical line that passes through the point  $(a, 0)$ . Every point on this line has the  $x$ -coordinate equal to  $a$ , regardless of the  $y$ -coordinate.

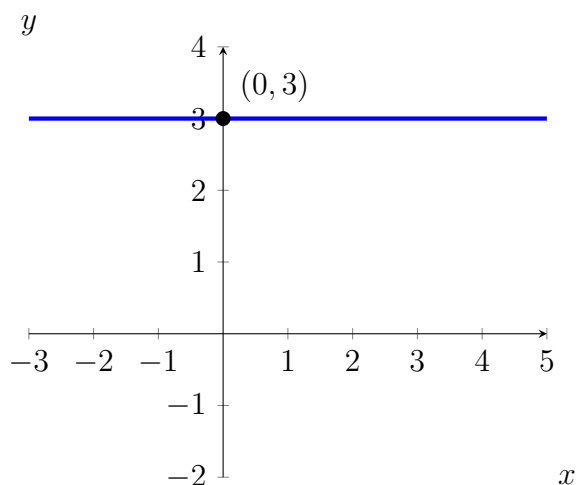
The graph of the line  $y = b$ , where  $b$  is a constant, is a horizontal line that passes through the point  $(0, b)$ . Every point on this line has the  $y$ -coordinate equal to  $b$ , regardless of the  $x$ -coordinate.

**Example 1.1.5.** *Graph the lines:  $x = -2$ , and  $y = 3$ .*

**Solution 1.1.5.** *The graph of the line  $x = -2$  is a vertical line that has the  $x$ -coordinate  $-2$  no matter what the  $y$ -coordinate is. The graph is a vertical line passing through point  $(-2, 0)$ .*



The graph of the line  $y = 3$  is a horizontal line that has the  $y$ -coordinate 3 regardless of what the  $x$ -coordinate is. Therefore, the graph is a horizontal line that passes through point  $(0, 3)$ .



## 1.2 Slope of a Line

In this section, you will learn to:

1. Find the slope of a line.
2. Graph the line if a point and the slope are given.

In the last section, we learned to graph a line by choosing two points on the line. A graph of a line can also be determined if one point and the "steepness" of the line is known. The number that refers to the steepness or inclination of a line is called the slope of the line. From previous math courses, many of you remember slope as the "rise over run," or "the vertical change over the horizontal change" and have often seen it expressed as:

$$\text{slope} = \frac{y_2 - y_1}{x_2 - x_1}$$

We give a precise definition.

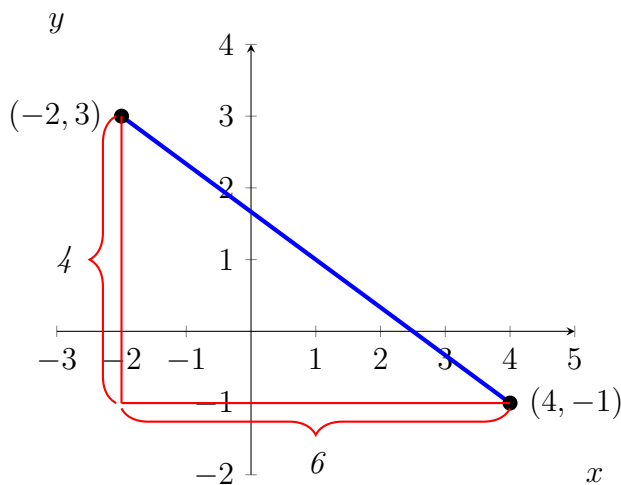
**Definition 1.2.1.** *If  $(x_1, y_1)$  and  $(x_2, y_2)$  are two different points on a line, the slope of the line is*

$$\text{slope} = m = \frac{y_2 - y_1}{x_2 - x_1}$$

**Example 1.2.1.** *Find the slope of the line passing through points  $(-2, 3)$  and  $(4, -1)$ , and graph the line.*

**Solution 1.2.1.** *Let  $(x_1, y_1) = (-2, 3)$  and  $(x_2, y_2) = (4, -1)$ , then the slope is*

$$\text{slope} = m = \frac{-1 - 3}{4 - (-2)} = \frac{-4}{6} = -\frac{2}{3}$$





To give the reader a better understanding, both the vertical change,  $-4$ , and the horizontal change,  $6$ , are shown in the above figure.

When two points are given, it does not matter which point is denoted as  $(x_1, y_1)$  and which  $(x_2, y_2)$ . The value for the slope will be the same.

In Example 1.2.1, if we instead choose  $(x_1, y_1) = (4, -1)$  and  $(x_2, y_2) = (-2, 3)$ , then we will get the same value for the slope as we obtained earlier.

The steps involved are as follows:

$$m = \frac{3 - (-1)}{-2 - 4} = \frac{4}{-6} = -\frac{2}{3}$$

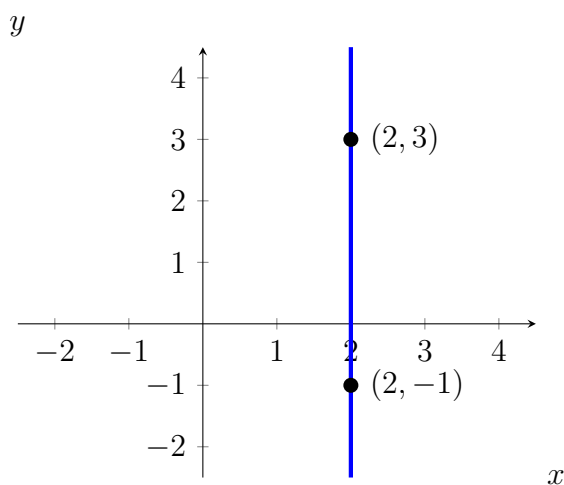
The student should further observe that

- If a line rises when going from left to right, then it has a positive slope. In this situation, as the value of  $x$  increases, the value of  $y$  also increases.
- If a line falls going from left to right, it has a negative slope; as the value of  $x$  increases, the value of  $y$  decreases.

**Example 1.2.2.** Find the slope of the line that passes through the points  $(2, 3)$  and  $(2, -1)$ , and graph.

**Solution 1.2.2.** Let  $(x_1, y_1) = (2, 3)$  and  $(x_2, y_2) = (2, -1)$ , then the slope is

$$m = \frac{-1 - 3}{2 - 2} = \frac{-4}{0} = \text{undefined}.$$

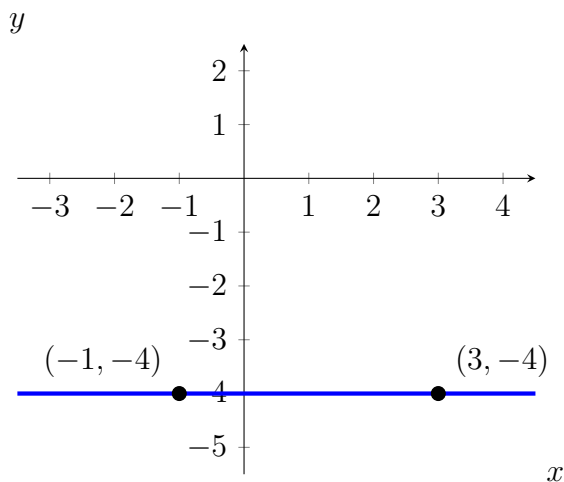


**Note 1.2.1.** *The slope of a vertical line is undefined.*

**Example 1.2.3.** *Find the slope of the line that passes through the points  $(-1, -4)$  and  $(3, -4)$ .*

**Solution 1.2.3.** *Let  $(x_1, y_1) = (-1, -4)$  and  $(x_2, y_2) = (3, -4)$ , then the slope is*

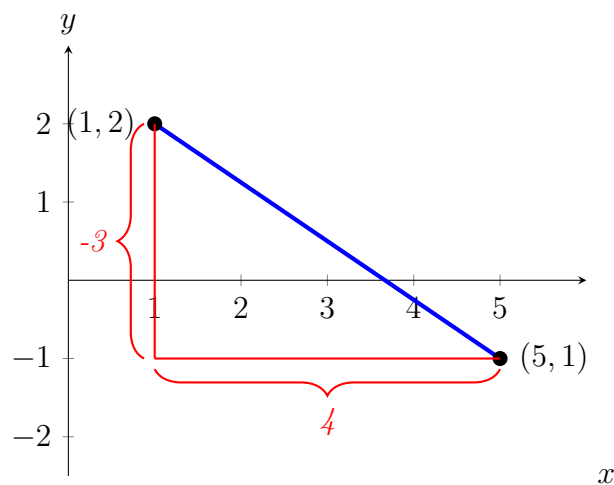
$$m = \frac{-4 - (-4)}{3 - (-1)} = \frac{0}{4} = 0.$$



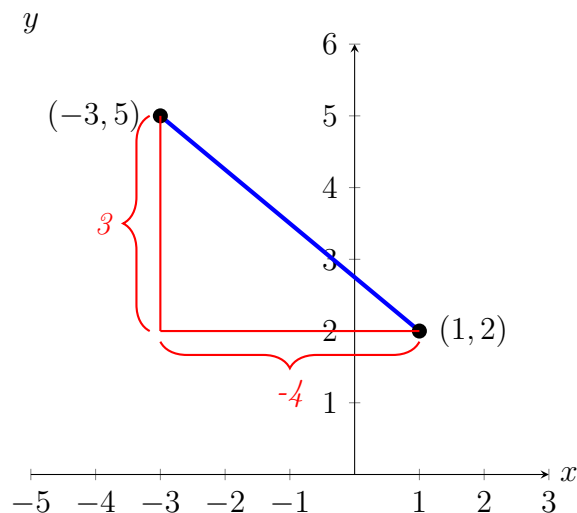
**Note 1.2.2.** *The slope of a horizontal line is 0.*

**Example 1.2.4.** *Graph the line that passes through the point  $(1, 2)$  and has a slope of  $-\frac{3}{4}$ .*

**Solution 1.2.4.** *The slope equals  $\frac{\text{rise}}{\text{run}}$ . The fact that the slope is  $-\frac{3}{4}$  means that for every rise of  $-3$  units (fall of 3 units), there is a run of 4 units. So if from the given point  $(1, 2)$  we go down 3 units and go right 4 units, we reach the point  $(5, -1)$ . The graph is obtained by connecting these two points.*



Alternatively, since  $\frac{3}{-4}$  represents the same number, the line can be drawn by starting at the point  $(1, 2)$  and choosing a rise of 3 units followed by a run of  $-4$  units. So from the point  $(1, 2)$ , we go up 3 units and to the left 4 units, thus reaching the point  $(-3, 5)$ , which is also on the same line. See figure below.



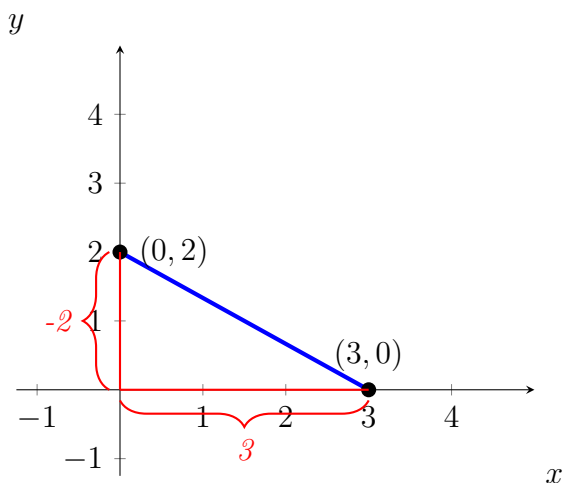
**Example 1.2.5.** Find the slope of the line  $2x + 3y = 6$ .

**Solution 1.2.5.** In order to find the slope of this line, we will choose any two points on this line. Again, the selection of  $x$  and  $y$  intercepts seems to be a good choice. The  $x$ -intercept is  $(3, 0)$ , and the  $y$ -intercept is  $(0, 2)$ . Therefore,

the slope is

$$m = \frac{2 - 0}{3 - 0} = \frac{-2}{3}.$$

The graph below shows the line and the  $x$ -intercepts and  $y$ -intercepts:



**Example 1.2.6.** Find the slope of the line  $y = 3x + 2$ .

**Solution 1.2.6.** We again find two points on the line, say  $(0, 2)$  and  $(1, 5)$ . Therefore, the slope is

$$m = \frac{5 - 2}{1 - 0} = \frac{3}{1} = 3.$$

Look at the slopes and the  $y$ -intercepts of the following lines.

Line	Slope	Y-Intercept
$y = 3x + 2$	3	2
$y = -2x + 5$	-2	5
$y = \frac{3}{2}x - 4$	$\frac{3}{2}$	-4

It is no coincidence that when an equation of the line is solved for  $y$ , the coefficient of the  $x$  term represents the slope, and the constant term represents the  $y$ -intercept. In other words, for the line  $y = mx + b$ ,  $m$  is the slope, and  $b$  is the  $y$ -intercept.

**Example 1.2.7.** Determine the slope and  $y$ -intercept of the line  $2x + 3y = 6$ .

**Solution 1.2.7.** We solve for  $y$ :

$$2x + 3y = 6$$

$$3y = -2x + 6$$

$$y = -\frac{2}{3}x + 2$$

*The slope is equal to the coefficient of the  $x$  term, which is  $-\frac{2}{3}$ . The  $y$ -intercept is equal to the constant term, which is 2.*

## 1.3 Determining the Equation of a Line

In this section, you will learn to:

1. Find an equation of a line if a point and the slope are given.
2. Find an equation of a line if two points are given.

So far, we were given an equation of a line and were asked to give information about it. For example, we were asked to find points on the line, find its slope, and even find intercepts. Now we are going to reverse the process. That is, we will be given either two points or a point and the slope of a line, and we will be asked to find its equation.

An equation of a line can be written in three forms: the slope-intercept form, the point-slope form, or the standard form. We will discuss each of them in this section.

A line is completely determined by two points or by a point and slope. The information we are given about a particular line will influence which form of the equation is most convenient to use. Once we know any form of the equation of a line, it is easy to re-express the equation in the other forms if needed.

**The Slope-Intercept Form of a Line:**  $y = mx + b$

In the last section, we learned that the equation of a line whose slope =  $m$  and  $y$ -intercept =  $b$  is  $y = mx + b$ . This is called the slope-intercept form of the line and is the most commonly used form.

**Example 1.3.1.** *Find an equation of a line whose slope is 5, and  $y$ -intercept is 3.*

**Solution 1.3.1.** *Since the slope is  $m = 5$ , and the  $y$ -intercept is  $b = 3$ , the equation is  $y = 5x + 3$ .*

**Example 1.3.2.** Find the equation of the line that passes through the point  $(2, 7)$  and has slope 3.

**Solution 1.3.2.** Since  $m = 3$ , the partial equation is  $y = 3x + b$ . Now  $b$  can be determined by substituting the point  $(2, 7)$  in the equation  $y = 3x + b$ .

$$7 = 3(2) + b$$

$$b = 1$$

Therefore, the equation is  $y = 3x + 1$ .

**Example 1.3.3.** Find an equation of the line that passes through the points  $(-1, 2)$  and  $(1, 8)$ .

**Solution 1.3.3.**  $m = \frac{8-2}{1-(-1)} = \frac{6}{2} = 3$ .

So the partial equation is  $y = 3x + b$ . We can use either of the two points  $(-1, 2)$  or  $(1, 8)$  to find  $b$ . Substituting  $(-1, 2)$  gives

$$2 = 3(-1) + b$$

$$5 = b$$

So the equation is  $y = 3x + 5$ .

**Example 1.3.4.** Find an equation of the line that has  $x$ -intercept 3, and  $y$ -intercept 4.

**Solution 1.3.4.** The  $x$ -intercept = 3, and  $y$ -intercept = 4 correspond to the points  $(3, 0)$  and  $(0, 4)$ , respectively.

$$m = \frac{4-0}{0-3} = \frac{-4}{3}$$

We are told the  $y$ -intercept is 4; thus  $b = 4$ .

Therefore, the equation is  $y = -\frac{4}{3}x + 4$ .

**The Point-Slope Form of a Line:**  $y - y_1 = m(x - x_1)$

The point-slope form is useful when we know two points on the line and want to find the equation of the line.

Let  $L$  be a line with slope  $m$ , and known to contain a specific point  $(x_1, y_1)$ . If  $(x, y)$  is any other point on the line  $L$ , then the definition of a slope leads us to the point-slope form or point-slope formula.

The slope is  $\frac{y-y_1}{x-x_1} = m$

Multiplying both sides by  $(x - x_1)$  gives the point-slope form:

$$y - y_1 = m(x - x_1)$$

**Example 1.3.5.** Find the point-slope form of the equation of a line that has slope 1.5 and passes through the point  $(12, 4)$ .

**Solution 1.3.5.** Substituting the point  $(x_1, y_1) = (12, 4)$  and  $m = 1.5$  in the point-slope formula, we get

$$y - 4 = 1.5(x - 12)$$

The student may be tempted to simplify this into the slope-intercept form  $y = mx + b$ . But since the problem specifically requests point-slope form, we will not simplify it.

**The Standard Form of a Line:**  $Ax + By = C$

Another useful form of the equation of a line is the standard form.

If we know the equation of a line in point-slope form,  $y - y_1 = m(x - x_1)$ , or if we know the equation of the line in slope-intercept form  $y = mx + b$ , we can simplify the formula to have all terms for the  $x$  and  $y$  variables on one side of the equation, and the constant on the other side of the equation.

The result is referred to as the standard form of the line:  $Ax + By = C$ .

**Example 1.3.6.** Using the point-slope formula, find the standard form of an equation of the line that passes through the point  $(2, 3)$  and has slope  $-\frac{3}{5}$ .

**Solution 1.3.6.** Substituting the point  $(2, 3)$  and  $m = -\frac{3}{5}$  in the point-slope formula, we get

$$y - 3 = -\frac{3}{5}(x - 2).$$

Multiplying both sides by 5 gives us

$$5(y - 3) = -3(x - 2),$$

$$5y - 15 = -3x + 6,$$

$$3x + 5y = 21 \text{ Standard Form.}$$

**Example 1.3.7.** Find the standard form of the line that passes through the points  $(1, -2)$  and  $(4, 0)$ .

**Solution 1.3.7.** First, we find the slope:  $m = \frac{0 - (-2)}{4 - 1} = \frac{2}{3}$ .

Then, the point-slope form is:  $y - (-2) = \frac{2}{3}(x - 1)$ .

Multiplying both sides by 3 gives us

$$3(y + 2) = 2(x - 1),$$

$$3y + 6 = 2x - 2,$$

$$-2x + 3y = -8,$$

$$2x - 3y = 8 \text{ Standard Form.}$$

**Example 1.3.8.** Write the equation  $y = -\frac{2}{3}x + 3$  in the standard form.

**Solution 1.3.8.** Multiplying both sides of the equation by 3, we get

$$3y = -2x + 9,$$

$$2x + 3y = 9 \text{ Standard Form.}$$

**Example 1.3.9.** Write the equation  $3x - 4y = 10$  in the slope-intercept form.

**Solution 1.3.9.** Solving for  $y$ , we get

$$-4y = -3x + 10,$$

$$y = \frac{3}{4}x - \frac{5}{2} \text{ Slope Intercept Form.}$$

**Example 1.3.10.** Find the slope of the following lines, by inspection.

1.  $3x - 5y = 10$

2.  $2x + 7y = 20$

3.  $4x - 3y = 8$

**Solution 1.3.10.** 1. For  $3x - 5y = 10$ , we have  $A = 3$  and  $B = -5$ , therefore,  $m = -\frac{A}{B} = -\frac{3}{-5} = \frac{3}{5}$ .

2. For  $2x + 7y = 20$ , we have  $A = 2$  and  $B = 7$ , therefore,  $m = -\frac{A}{B} = -\frac{2}{7}$ .



3. For  $4x - 3y = 8$ , we have  $A = 4$  and  $B = -3$ , therefore,  $m = -\frac{A}{B} = -\frac{4}{-3} = \frac{4}{3}$ .

**Example 1.3.11.** Find an equation of the line that passes through  $(2, 3)$  and has slope  $-\frac{4}{5}$ .

**Solution 1.3.11.** Since the slope of the line is  $-\frac{4}{5}$ , we know that the left side of the equation is  $4x + 5y$ , and the partial equation is going to be

$$4x + 5y = c.$$

Of course,  $c$  can easily be found by substituting for  $x$  and  $y$ .

$$4(2) + 5(3) = c,$$

$$8 + 15 = c,$$

$$23 = c.$$

The desired equation is

$$4x + 5y = 23.$$

If you use this method often enough, you can do these problems very quickly. We summarize the forms for equations of a line below:

#### Summary 1.3.1: E

equations of Lines

- Slope-Intercept form:  $y = mx + b$ , where  $m$  is the slope and  $b$  is the  $y$ -intercept.
- Point-Slope form:  $y - y_1 = m(x - x_1)$ , where  $m$  is the slope and  $(x_1, y_1)$  is a point on the line.
- Standard form:  $Ax + By = C$ .
- Horizontal Line:  $y = b$ , where  $b$  is the  $y$ -intercept.
- Vertical Line:  $x = a$ , where  $a$  is the  $x$ -intercept.

## 1.4 Applications

In this section, you will learn to use linear functions to model real-world applications.

Now that we have learned to determine equations of lines, we get to apply these ideas in a variety of real-life situations. Read the problem carefully. Highlight important information. Keep track of which values correspond to the independent variable ( $x$ ) and which correspond to the dependent variable ( $y$ ).

**Example 1.4.1.** *A taxi service charges \$0.50 per mile plus a \$5 flat fee. What will be the cost of traveling 20 miles? What will be cost of traveling  $x$  miles?*

**Solution 1.4.1.** *Let  $x$  be the distance traveled, in miles, and  $y$  be the cost in dollars.*

*The cost of traveling 20 miles is  $y = (0.50)(20) + 5 = 10 + 5 = 15$  dollars.*

*The cost of traveling  $x$  miles is  $y = (0.50)(x) + 5 = 0.50x + 5$  dollars.*

*In this problem, \$0.50 per mile is referred to as the variable cost, and the flat charge \$5 as the fixed cost. Now if we look at our cost equation  $y = 0.50x + 5$ , we can see that the variable cost corresponds to the slope and the fixed cost to the  $y$ -intercept.*

**Example 1.4.2.** *The variable cost to manufacture a product is \$10 per item and the fixed cost \$2500. If  $x$  represents the number of items manufactured and  $y$  represents the total cost, write the cost function.*

**Solution 1.4.2.** *The variable cost of \$10 per item tells us that  $m = 10$ . The fixed cost represents the  $y$ -intercept, so  $b = 2500$ . Therefore, the cost equation is  $y = 10x + 2500$ .*

**Example 1.4.3.** *It costs \$750 to manufacture 25 items, and \$1000 to manufacture 50 items. Assuming a linear relationship holds, find the cost equation, and use this function to predict the cost of 100 items.*

**Solution 1.4.3.** *Let  $x$  be the number of items manufactured, and let  $y$  be the cost.*

*Solving this problem is equivalent to finding an equation of a line that passes through the points (25, 750) and (50, 1000).*

$$m = \frac{1000-750}{50-25} = 10$$

*Therefore, the partial equation is  $y = 10x + b$ .*

By substituting one of the points in the equation, we get  $b = 500$ .

Therefore, the cost equation is  $y = 10x + 500$ .

To find the cost of 100 items, substitute  $x = 100$  in the equation  $y = 10x + 500$ .

So the cost  $= y = 10(100) + 500 = 1500$ .

It costs \$1500 to manufacture 100 items.

**Example 1.4.4.** The freezing temperature of water in Celsius is 0 degrees, and in Fahrenheit, it's 32 degrees. The boiling temperatures of water in Celsius and Fahrenheit are 100 degrees and 212 degrees, respectively. Write a conversion equation from Celsius to Fahrenheit and use this equation to convert 30 degrees Celsius into Fahrenheit.

**Solution 1.4.4.** Let's look at what is given:

Celsius	Fahrenheit
0	32
100	212

Solving this problem is equivalent to finding an equation of a line that passes through the points  $(0, 32)$  and  $(100, 212)$ . Since we are finding a linear relationship, we are looking for an equation  $y = mx + b$ , or in this case,  $F = mC + b$ , where  $C$  represents the temperature in Celsius, and  $F$  represents the temperature in Fahrenheit.

The slope  $m = \frac{212-32}{100-0} = 95$ .

The equation is  $F = 95C + b$ .

Substituting the point  $(0, 32)$ , we get  $F = 95C + 32$ .

To convert 30 degrees Celsius into Fahrenheit, substitute  $C = 30$  in the equation:

$$F = 95C + 32$$

$$F = 95(30) + 32 = 86$$

**Example 1.4.5.** The population of Canada in the year 1980 was 24.5 million, and in the year 2010, it was 34 million. The population of Canada over that time period can be approximately modeled by a linear function. Let  $x$

represent time as the number of years after 1980, and let  $y$  represent the size of the population.

- a. Write the linear function that gives a relationship between the time and the population.
- b. Assuming the population continues to grow linearly in the future, use this equation to predict the population of Canada in the year 2025.

**Solution 1.4.5.** The problem can be made easier by using 1980 as the base year, which means we choose the year 1980 as the year zero. This will make the year 2010 correspond to year 30. Now, let's look at the information we have:

Year	Population
0 (1980)	24.5 million
30 (2010)	34 million

- a. Solving this problem is equivalent to finding an equation of a line that passes through the points  $(0, 24.5)$  and  $(30, 34)$ . We use these two points to find the slope:

$$m = \frac{34 - 24.5}{30 - 0} = \frac{9.5}{30} = 0.32$$

The  $y$ -intercept occurs when  $x = 0$ , so  $b = 24.5$ .

So, the equation relating time ( $x$ ) and population ( $y$ ) is:

$$y = 0.32x + 24.5$$

- b. Now, to predict the population in the year 2025, we let  $x = 2025 - 1980 = 45$ :

$$y = 0.32x + 24.5$$

$$y = 0.32(45) + 24.5 = 38.9$$

In the year 2025, we predict that the population of Canada will be 38.9 million people.

*Note that we assumed the population trend will continue to be linear. Therefore, if population trends change and this assumption does not continue to be true in the future, this prediction may not be accurate.*

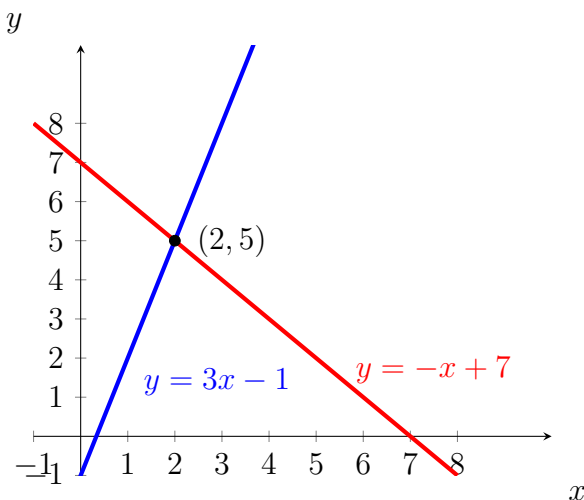
## 1.5 More Applications

### 1.5.1 Finding the Point of Intersection of Two Lines

In this section, we will do application problems that involve the intersection of lines. Therefore, before we proceed any further, we will first learn how to find the intersection of two lines.

**Example 1.5.1.** *Find the intersection of the line  $y = 3x - 1$  and the line  $y = -x + 7$ .*

**Solution 1.5.1.** *We graph both lines on the same axes, as shown below, and read the solution  $(2, 5)$ .*



*Finding an intersection of two lines graphically is not always easy or practical; therefore, we will now learn to solve these problems algebraically.*

*At the point where two lines intersect, the  $x$  and  $y$  values for both lines are the same. So in order to find the intersection, we either let the  $x$ -values or the  $y$ -values equal.*

*If we were to solve the above example algebraically, it will be easier to let the*

*y*-values equal. Since  $y = 3x - 1$  for the first line, and  $y = -x + 7$  for the second line, by letting the *y*-values equal, we get:

$$3x - 1 = -x + 7$$

$$4x = 8$$

$$x = 2$$

By substituting  $x = 2$  in any of the two equations, we obtain  $y = 5$ . Hence, the solution is  $(2, 5)$ .

### 1.5.2 Solving Systems of Equations: The Elimination Method

A common algebraic method used to solve systems of equations is called the elimination method. The objective is to eliminate one of the two variables by adding the left and right sides of the equations together. Once one variable is eliminated, we have an equation with only one variable that can be solved. Finally, by substituting the value of the variable that has been found in one of the original equations, we can get the value of the other variable.

**Example 1.5.2.** Find the intersection of the lines  $2x + y = 7$  and  $3x - y = 3$  by the elimination method.

**Solution 1.5.2.** We add the left and right sides of the two equations:

$$2x + y = 7$$

$$3x - y = 3$$

$$5x = 10$$

$$x = 2$$

Now we substitute  $x = 2$  into any of the two equations and solve for  $y$ :

$$\begin{aligned}2(2) + y &= 7 \\4 + y &= 7 \\y &= 3\end{aligned}$$

Therefore, the solution is  $(2, 3)$ .

**Example 1.5.3.** Solve the system of equations  $x + 2y = 3$  and  $2x + 3y = 4$  by the elimination method.

**Solution 1.5.3.** If we add the two equations directly, none of the variables are eliminated. However, the variable  $x$  can be eliminated by multiplying the first equation by  $-2$  and leaving the second equation unchanged:

$$\begin{aligned}-2x - 4y &= -6 \\2x + 3y &= 4 \\-y &= -2 \\y &= 2\end{aligned}$$

Substituting  $y = 2$  into  $x + 2y = 3$ , we get:

$$\begin{aligned}x + 2(2) &= 3 \\x + 4 &= 3 \\x &= -1\end{aligned}$$

Therefore, the solution is  $(-1, 2)$ .

**Example 1.5.4.** Solve the system of equations  $3x - 4y = 5$  and  $4x - 5y = 6$ .

**Solution 1.5.4.** This time, we multiply the first equation by  $-4$  and the second by  $3$  before adding (the choice of numbers is not unique):

$$\begin{aligned}-12x + 16y &= -20 \\12x - 15y &= 18 \\y &= -2\end{aligned}$$

By substituting  $y = -2$  into any one of the equations, we get:

$$3x - 4(-2) = 5$$

$$3x + 8 = 5$$

$$3x = -3$$

$$x = -1$$

Hence, the solution is  $(-1, -2)$ .

### 1.5.3 Supply, Demand, and the Equilibrium Market Price

In a free market economy, the supply curve for a commodity is the number of items of a product that can be made available at different prices, and the demand curve is the number of items the consumer will buy at different prices. As the price of a product increases, its demand decreases, and supply increases. On the other hand, as the price decreases, the demand increases, and supply decreases. The equilibrium price is reached when the demand equals the supply.

**Example 1.5.5.** *The supply curve for a product is given by  $y = 3.5x - 14$ , and the demand curve for the same product is given by  $y = -2.5x + 34$ , where  $x$  is the price and  $y$  is the number of items produced. Find the following:*

1. *How many items will be supplied at a price of \$10?*
2. *How many items will be demanded at a price of \$10?*
3. *Determine the equilibrium price.*
4. *How many items will be produced at the equilibrium price?*

**Solution 1.5.5.** 1. *To find the number of items supplied at a price of \$10, we substitute  $x = 10$  into the supply equation  $y = 3.5x - 14$ . Therefore,  $y = 3.5(10) - 14 = 21$  items will be supplied.*

2. *To find the number of items demanded at a price of \$10, we substitute  $x = 10$  into the demand equation  $y = -2.5x + 34$ . Therefore,  $y = -2.5(10) + 34 = 9$  items will be demanded.*



3. To determine the equilibrium price, we set the supply equal to the demand:

$$3.5x - 14 = -2.5x + 34$$

Solving for  $x$ :

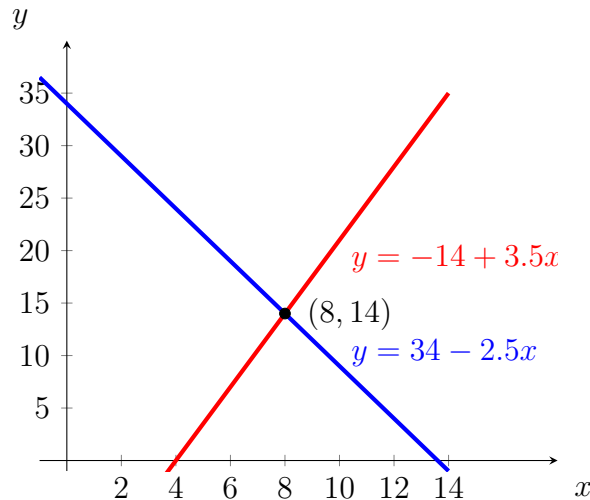
$$6x = 48$$

$$x = 8$$

So, the equilibrium price is  $x = 8$ .

4. To find how many items will be produced at the equilibrium price, we substitute  $x = 8$  into either the supply or the demand equation. Using the supply equation, we get  $y = 3.5(8) - 14 = 14$  items will be produced.

The graph shows the intersection of the supply and demand functions and their point of intersection,  $(8, 14)$ .



#### 1.5.4 Break-Even Point

In a business, profit is generated by selling products. If a company sells  $x$  number of items at a price  $P$ , then the revenue  $R$  is the price multiplied by the number of items sold:  $R = P \cdot x$ . The production costs  $C$  are the sum of the variable costs and the fixed costs, often written as  $C = mx + b$ , where  $x$  is the number of items manufactured.

- The slope  $m$  is called the marginal cost and represents the cost to produce one additional item or unit.
- The variable cost,  $mx$ , depends on how much is being produced.
- The fixed cost  $b$  is constant and does not change regardless of production quantity.

Profit is equal to revenue minus cost:  $Profit = R - C$ . A company makes a profit if the revenue is greater than the cost, and there is a loss if the cost is greater than the revenue. The point on the graph where the revenue equals the cost is called the break-even point, and at this point, the profit is 0.

**Example 1.5.6.** *If the revenue function of a product is  $R = 5x$  and the cost function is  $C = 3x + 12$ , find the following:*

1. *If 4 items are produced, what will the revenue be?*
2. *What is the cost of producing 4 items?*
3. *How many items should be produced to break even?*
4. *What will be the revenue and cost at the break-even point?*

**Solution 1.5.6.** 1. *To find the revenue when 4 items are produced, we substitute  $x = 4$  in the revenue equation  $R = 5x$ , and the answer is  $R = 20$ .*

2. *To find the cost of producing 4 items, we substitute  $x = 4$  in the cost equation  $C = 3x + 12$ , and the answer is  $C = 24$ .*
3. *To determine the number of items required to break even, we set the revenue equal to the cost:*

$$5x = 3x + 12$$

*Solving for  $x$ :*

$$2x = 12$$

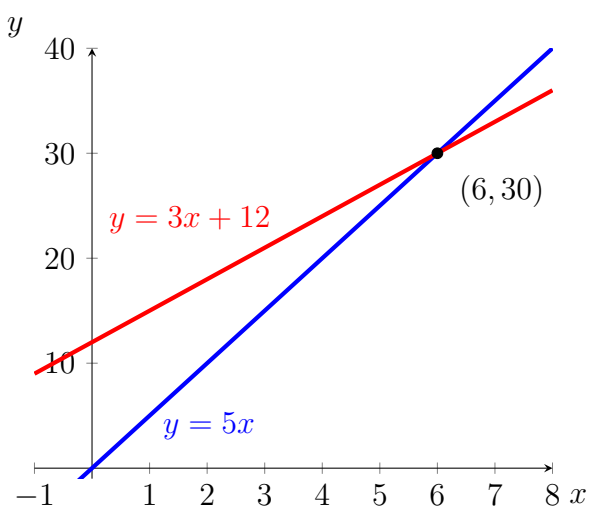
$$x = 6$$

*So, 6 items should be produced to break even.*

4. *At the break-even point, when  $x = 6$ , we can substitute  $x = 6$  in either the revenue or the cost equation to find that both revenue and cost are*

equal to 30. Therefore, the revenue and cost at the break-even point are both 30.

The graph below shows the intersection of the revenue and cost functions and their point of intersection,  $(6, 30)$ .





# Chapter 2

## Matrices

In this chapter, you will learn to:

1. Do matrix operations.
2. Solve linear systems using the Gauss-Jordan method.
3. Solve linear systems using the matrix inverse method.
4. Do application problems

### 2.1 Introduction to Matrices

In this section, you will learn to:

1. Add and subtract matrices.
2. Multiply a matrix by a scalar.
3. Multiply two matrices.

A matrix is a 2-dimensional array of numbers arranged in rows and columns. Matrices provide a method of organizing, storing, and working with mathematical information. They have numerous applications and uses in the real world.

(TODO: fix references here)Matrices are particularly useful in working with models based on systems of linear equations, which we'll explore in sections

2.2, 2.3, and 2.4 of this chapter. They are also used in encryption (section 2.5) and economic modeling (section 2.6).

Furthermore, matrices play a crucial role in optimization problems in (TODO: fix references here) Chapter 4, such as maximizing profit or revenue and minimizing costs. They are used in business for scheduling, routing transportation and shipments, and managing inventory. Matrices are applicable in various fields where data organization and problem-solving are essential.

The use of matrices has expanded with the increase in available data across different domains. They are fundamental tools for organizing data and solving problems in science fields like physics, chemistry, biology, genetics, meteorology, and economics. In computer science, matrix mathematics is foundational for animation in movies and video games.

Moreover, matrices are used in analyzing network diagrams, such as social media connections on platforms like Facebook, LinkedIn, etc. The mathematics of network diagrams falls under "graph theory" and relies on matrices to organize information in graphs that depict connections and associations in a network.

A matrix is a rectangular array of numbers. Matrices are useful in organizing and manipulating large amounts of data. In order to get some idea of what matrices are all about, we will look at the following example.

**Example 2.1.1.** *Fine Furniture Company makes chairs and tables at its San Jose, Hayward, and Oakland factories. The total production, in hundreds, from the three factories for the years 2014 and 2015 is listed in the table below.*

	2014	2015		
	CHAIRS	TABLES	CHAIRS	TABLES
SAN JOSE	30	18	36	20
HAYWARD	20	12	24	18
OAKLAND	16	10	20	12

1. Represent the production for the years 2014 and 2015 as the matrices  $A$  and  $B$ .
2. Find the difference in sales between the years 2014 and 2015.

3. The company predicts that in the year 2020 the production at these factories will be double that of the year 2014. What will the production be for the year 2020?

**Solution 2.1.1.** 1. The matrices are as follows:

$$A = \begin{bmatrix} 30 & 18 \\ 20 & 12 \\ 16 & 10 \end{bmatrix} \quad B = \begin{bmatrix} 36 & 20 \\ 24 & 18 \\ 20 & 12 \end{bmatrix}$$

2. We are looking for the matrix  $B - A$ . When two matrices have the same number of rows and columns, they can be added or subtracted entry by entry. Therefore, we get:

$$B - A = \begin{bmatrix} 36 - 30 & 20 - 18 \\ 24 - 20 & 18 - 12 \\ 20 - 16 & 12 - 10 \end{bmatrix} = \begin{bmatrix} 6 & 2 \\ 4 & 6 \\ 4 & 2 \end{bmatrix}$$

3. We would like a matrix that is twice the matrix of 2014, i.e.,  $2A$ . Whenever a matrix is multiplied by a number, each entry is multiplied by the number.

$$2A = 2 \begin{bmatrix} 30 & 18 \\ 20 & 12 \\ 16 & 10 \end{bmatrix} = \begin{bmatrix} 60 & 36 \\ 40 & 24 \\ 32 & 20 \end{bmatrix}$$

### 2.1.1 Vocabulary

Before we go any further, we need to familiarize ourselves with some terms that are associated with matrices.

The numbers in a matrix are called the entries or the elements of a matrix.

Whenever we talk about a matrix, we need to know its size or dimension. The dimension of a matrix is the number of rows and columns it has. When we say a matrix is a "3 by 4 matrix," we are saying that it has 3 rows and 4 columns. The rows are always mentioned first, and the columns second. This means that a  $3 \times 4$  matrix does not have the same dimension as a  $4 \times 3$  matrix.

$$A = \begin{bmatrix} 1 & 4 & -2 & 0 \\ 3 & -1 & 7 & 9 \\ 6 & 2 & 0 & 5 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 9 & 8 \\ -3 & 0 & 1 \\ 6 & 5 & -2 \\ -4 & 7 & 8 \end{bmatrix}$$

Matrix  $A$  has dimensions  $3 \times 4$  — Matrix  $B$  has dimensions  $4 \times 3$

A matrix that has the same number of rows as columns is called a square matrix. A matrix with all entries zero is called a zero matrix. A square matrix with 1's along the main diagonal and zeros everywhere else, is called an identity matrix. When a square matrix is multiplied by an identity matrix of same size, the matrix remains the same.

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Matrix  $I$  is a  $3 \times 3$  identity matrix

A matrix with only one row is called a row matrix or a row vector, and a matrix with only one column is called a column matrix or a column vector. Two matrices are equal if they have the same size and the corresponding entries are equal. We can perform arithmetic operations with matrices. Next we will define and give examples illustrating the operations of matrix addition and subtraction, scalar multiplication, and matrix multiplication. Note that matrix multiplication is quite different from what you would intuitively expect, so pay careful attention to the explanation. Note also that the ability to perform matrix operations depends on the matrices involved being compatible in size, or dimensions, for that operation. The definition of compatible dimensions is different for different operations, so note the requirements carefully for each.

### 2.1.2 Matrix Addition and Subtraction

If two matrices have the same size, they can be added or subtracted. The operations are performed on corresponding entries.



**Example 2.1.2.** *Given the matrices  $A$ ,  $B$ ,  $C$ , and  $D$  below:*

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 3 & 1 \\ 5 & 0 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -1 & 3 \\ 2 & 4 & 2 \\ 3 & 6 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix}, \quad D = \begin{bmatrix} -2 \\ -3 \\ 4 \end{bmatrix}$$

*Find, if possible:*

1.  $A + B$
2.  $C - D$
3.  $A + D$

**Solution 2.1.2.**     • *We add each element of  $A$  to the corresponding entry of  $B$ :*

$$A + B = \begin{bmatrix} 3 & 1 & 7 \\ 4 & 7 & 3 \\ 8 & 6 & 4 \end{bmatrix}$$

- *We perform the subtraction entry by entry for  $C - D$ :*

$$C - D = \begin{bmatrix} 6 \\ 5 \\ -1 \end{bmatrix}$$

- *The sum  $A + D$  cannot be found because the two matrices have different sizes. Two matrices can only be added or subtracted if they have the same dimension.*

### 2.1.3 Multiplying a Matrix by a Scalar

If a matrix is multiplied by a scalar, each entry is multiplied by that scalar.

**Example 2.1.3.** *Given the matrix  $A$  and  $C$  in the previous example, find  $2A$  and  $-3C$ .*

**Solution 2.1.3.**     • *To find  $2A$ , we multiply each entry of matrix  $A$  by 2:*

$$2A = \begin{bmatrix} 2 & 4 & 8 \\ 4 & 6 & 2 \\ 10 & 0 & 6 \end{bmatrix}$$

- To find  $-3C$ , we multiply each entry of  $C$  by  $-3$ :

$$-3C = \begin{bmatrix} -12 \\ -6 \\ -9 \end{bmatrix}$$

### 2.1.4 Multiplication of Two Matrices

To multiply a matrix by another is not as easy as the addition, subtraction, or scalar multiplication of matrices. Because of its wide use in application problems, it is important that we learn it well. Therefore, we will try to learn the process in a step by step manner.

**Example 2.1.4.** Given  $A = \begin{bmatrix} 2 & 3 & 4 \end{bmatrix}$  and  $B = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ , find the product  $AB$ .

**Solution 2.1.4.** The product is a  $1 \times 1$  matrix whose entry is obtained by multiplying the corresponding entries and then forming the sum:

$$AB = \begin{bmatrix} 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 2a + 3b + 4c$$

Note that  $AB$  is a  $1 \times 1$  matrix, and its only entry is  $2a + 3b + 4c$ .

**Example 2.1.5.** Given  $A = \begin{bmatrix} 2 & 3 & 4 \end{bmatrix}$  and  $B = \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}$ , find the product  $AB$ .

**Solution 2.1.5.** Again, we multiply the corresponding entries and add:

$$AB = \begin{bmatrix} 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix} = (2 \cdot 5) + (3 \cdot 6) + (4 \cdot 7) = 10 + 18 + 28 = 56$$

**Example 2.1.6.** Given  $A = \begin{bmatrix} 2 & 3 & 4 \end{bmatrix}$  and  $B = \begin{bmatrix} 5 & 3 \\ 6 & 4 \\ 7 & 5 \end{bmatrix}$ , find the product  $AB$ .

**Solution 2.1.6.** We know how to multiply a row matrix by a column matrix. To find the product  $AB$ , in this example, we will multiply the row matrix  $A$  to both the first and second columns of matrix  $B$ , resulting in a  $1 \times 2$  matrix:

$$AB = [2 \quad 3 \quad 4] \begin{bmatrix} 5 & 3 \\ 6 & 4 \\ 7 & 5 \end{bmatrix} = [(2 \cdot 5) + (3 \cdot 6) + (4 \cdot 7) \quad (2 \cdot 3) + (3 \cdot 4) + (4 \cdot 5)] = [56 \quad 38]$$

We multiplied a  $1 \times 3$  matrix by a matrix whose size is  $3 \times 2$ . So unlike addition and subtraction, it is possible to multiply two matrices with different dimensions if the number of entries in the rows of the first matrix is the same as the number of entries in the columns of the second matrix.

**Example 2.1.7.** Given  $A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 3 \end{bmatrix}$  and  $B = \begin{bmatrix} 5 & 3 \\ 6 & 4 \\ 7 & 5 \end{bmatrix}$ , find the product  $AB$ .

**Solution 2.1.7.** This time we are multiplying two rows of matrix  $A$  with two columns of matrix  $B$ . Since the number of entries in each row of  $A$  is the same as the number of entries in each column of  $B$ , the product is possible. We do exactly what we did in the last example. The only difference is that matrix  $A$  has one more row.

We multiply the first row of matrix  $A$  with the two columns of  $B$ , one at a time, and then repeat the process with the second row of  $A$ . We get:

$$AB = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 5 & 3 \\ 6 & 4 \\ 7 & 5 \end{bmatrix} = \begin{bmatrix} (2 \cdot 5 + 3 \cdot 6 + 4 \cdot 7) & (2 \cdot 3 + 3 \cdot 4 + 4 \cdot 5) \\ (1 \cdot 5 + 2 \cdot 6 + 3 \cdot 7) & (1 \cdot 3 + 2 \cdot 4 + 3 \cdot 5) \end{bmatrix}$$

$$AB = \begin{bmatrix} 56 & 38 \\ 38 & 26 \end{bmatrix}$$

**Example 2.1.8.** Given matrices  $E = \begin{bmatrix} 1 & 2 \\ 4 & 2 \\ 3 & 1 \end{bmatrix}$ ,  $F = \begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix}$ ,  $G = [4 \quad 1]$ ,

and  $H = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$ , find the following products if possible:

1.  $EF$

2.  $FE$

3.  $FH$

4.  $GH$

5.  $HG$

**Solution 2.1.8.** 1. To find  $EF$ , we multiply the rows of  $E$  with the columns of  $F$ . The result is:

$$EF = \begin{bmatrix} 1 & 2 \\ 4 & 2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} (1 \cdot 2 + 2 \cdot 3) & (1 \cdot -1 + 2 \cdot 2) \\ (4 \cdot 2 + 2 \cdot 3) & (4 \cdot -1 + 2 \cdot 2) \\ (3 \cdot 2 + 1 \cdot 3) & (3 \cdot -1 + 1 \cdot 2) \end{bmatrix} = \begin{bmatrix} 8 & 3 \\ 14 & 0 \\ 9 & -1 \end{bmatrix}$$

2. Product  $FE$  is not possible because  $F$  has two entries in each row, while  $E$  has three entries in each column.

$$3. FH = \begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \end{bmatrix} = \begin{bmatrix} (2 \cdot -3 + -1 \cdot -1) \\ (3 \cdot -3 + 2 \cdot -1) \end{bmatrix} = \begin{bmatrix} -5 \\ -11 \end{bmatrix}$$

$$4. GH = [4 \quad 1] \begin{bmatrix} -3 \\ -1 \end{bmatrix} = (4 \cdot -3 + 1 \cdot -1) = -13$$

$$5. HG = \begin{bmatrix} -3 \\ -1 \end{bmatrix} [4 \quad 1] = \begin{bmatrix} (-3 \cdot 4 & -3 \cdot 1) \\ (-1 \cdot 4 & -1 \cdot 1) \end{bmatrix} = \begin{bmatrix} -12 & -3 \\ -4 & -1 \end{bmatrix}$$

We summarize some important properties of matrix multiplication that we observed in the previous examples.

- For the product  $AB$  to exist, the number of columns of matrix  $A$  must equal the number of rows of matrix  $B$ .
- If matrix  $A$  has dimensions  $m \times n$  and matrix  $B$  has dimensions  $n \times p$ , then the product  $AB$  will have dimensions  $m \times p$ .
- Matrix multiplication is not commutative; that is, in general,  $AB$  does not equal  $BA$ .

**Example 2.1.9.** Given matrices  $R = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 5 \\ 2 & 3 & 1 \end{bmatrix}$ ,  $S = \begin{bmatrix} 0 & -1 & 2 \\ 3 & 1 & 0 \\ 4 & 2 & 1 \end{bmatrix}$ , and

$$T = \begin{bmatrix} -2 & 3 & 0 \\ -3 & 2 & 2 \\ -1 & 1 & 0 \end{bmatrix}, \text{ find } 2RS - 3ST.$$

**Solution 2.1.9.** *Solution:* To find  $2RS - 3ST$ , we first compute the products  $RS$  and  $ST$ :

$$RS = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 5 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 & 2 \\ 3 & 1 & 0 \\ 4 & 2 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} (1 \cdot 0 + 0 \cdot 3 + 2 \cdot 4) & (1 \cdot -1 + 0 \cdot 1 + 2 \cdot 2) & (1 \cdot 2 + 0 \cdot 0 + 2 \cdot 1) \\ (2 \cdot 0 + 1 \cdot 3 + 5 \cdot 4) & (2 \cdot -1 + 1 \cdot 1 + 5 \cdot 2) & (2 \cdot 2 + 1 \cdot 0 + 5 \cdot 1) \\ (2 \cdot 0 + 3 \cdot 3 + 1 \cdot 4) & (2 \cdot -1 + 3 \cdot 1 + 1 \cdot 2) & (2 \cdot 2 + 3 \cdot 0 + 1 \cdot 1) \end{bmatrix}$$

$$= \begin{bmatrix} 8 & 3 & 4 \\ 23 & 9 & 9 \\ 13 & 3 & 5 \end{bmatrix}$$

Next, we compute  $ST$ :

$$ST = \begin{bmatrix} 0 & -1 & 2 \\ 3 & 1 & 0 \\ 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} -2 & 3 & 0 \\ -3 & 2 & 2 \\ -1 & 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} (0 \cdot -2 + -1 \cdot -3 + 2 \cdot -1) & (0 \cdot 3 + -1 \cdot 2 + 2 \cdot 1) & (0 \cdot 0 + -1 \cdot 2 + 2 \cdot 0) \\ (3 \cdot -2 + 1 \cdot -3 + 0 \cdot -1) & (3 \cdot 3 + 1 \cdot 2 + 0 \cdot 1) & (3 \cdot 0 + 1 \cdot 1 + 0 \cdot 0) \\ (4 \cdot -2 + 2 \cdot -3 + 1 \cdot -1) & (4 \cdot 3 + 2 \cdot 2 + 1 \cdot 1) & (4 \cdot 0 + 2 \cdot 1 + 1 \cdot 0) \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & -2 \\ -9 & 11 & 2 \\ -15 & 17 & 4 \end{bmatrix}$$

Now we can find  $2RS - 3ST$ :

$$2RS - 3ST = 2 \cdot \begin{bmatrix} 8 & 3 & 4 \\ 23 & 9 & 9 \\ 13 & 3 & 5 \end{bmatrix} - 3 \cdot \begin{bmatrix} 1 & 0 & -2 \\ -9 & 11 & 2 \\ -15 & 17 & 4 \end{bmatrix}$$

$$\begin{aligned}
&= \begin{bmatrix} 16 & 6 & 8 \\ 46 & 18 & 18 \\ 26 & 6 & 10 \end{bmatrix} - \begin{bmatrix} 3 & 0 & 6 \\ -27 & 33 & 6 \\ -45 & 51 & 12 \end{bmatrix} \\
&= \begin{bmatrix} 13 & 6 & 14 \\ 73 & -15 & 12 \\ 71 & -45 & -2 \end{bmatrix}
\end{aligned}$$

The result of  $2RS - 3ST$  is a matrix with dimensions  $3 \times 3$ .

**Example 2.1.10.** Given matrix  $F = \begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix}$ , find  $F^2$ .

**Solution 2.1.10.**  $F^2$  is found by multiplying matrix  $F$  by itself, using matrix multiplication.

$$F^2 = \begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 2 \cdot 2 + (-1) \cdot 3 & 2 \cdot (-1) + (-1) \cdot 2 \\ 3 \cdot 2 + 2 \cdot 3 & 3 \cdot (-1) + 2 \cdot 2 \end{bmatrix} = \begin{bmatrix} 1 & -4 \\ 12 & 1 \end{bmatrix}$$

Note that  $F^2$  is not found by squaring each entry of matrix  $F$ . The process of raising a matrix to a power, such as finding  $F^2$ , is only possible if the matrix is a square matrix.

### 2.1.5 Systems of Linear Equations

Using matrices to represent a system of linear equations is a powerful technique that allows for efficient solving of such systems. In this method, we define matrices as follows:

- Matrix  $A$  represents the coefficients of the variables in the system and is called the coefficient matrix.
- Matrix  $X$  is a column matrix that contains the variables of the system.
- Matrix  $B$  is a column matrix that contains the constants of the system.

By defining these matrices, we can represent a system of linear equations as the matrix equation  $AX = B$ , where  $A$ ,  $X$ , and  $B$  are matrices. This representation simplifies the process of solving linear systems and allows us to apply matrix operations to find the solution.

In the next sections, we will delve deeper into how to use matrices to solve linear systems and explore various methods and techniques for efficient computation and analysis. Matrix representation is widely used in mathematical modeling, engineering, economics, and various other fields where systems of linear equations arise.

**Example 2.1.11.** *Verify that the system of two linear equations with two unknowns:*

$$\begin{aligned} ax + by &= h \\ cx + dy &= k \end{aligned}$$

*can be written as  $AX = B$ , where*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad X = \begin{bmatrix} x \\ y \end{bmatrix}, \quad B = \begin{bmatrix} h \\ k \end{bmatrix}.$$

**Solution 2.1.11.** *If we multiply the matrices  $A$  and  $X$ , we get*

$$AX = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}.$$

*If  $AX = B$ , then*

$$\begin{bmatrix} ax + by \\ cx + dy \end{bmatrix} = \begin{bmatrix} h \\ k \end{bmatrix}.$$

*If two matrices are equal, then their corresponding entries are equal. It follows that*

$$\begin{aligned} ax + by &= h \\ cx + dy &= k \end{aligned}$$

**Example 2.1.12.** *Express the following system as a matrix equation in the form  $AX = B$ .*

$$\begin{aligned} 2x + 3y - 4z &= 5 \\ 3x + 4y - 5z &= 6 \\ 5x - 6z &= 7 \end{aligned}$$

**Solution 2.1.12.** *This system of equations can be expressed in the form  $AX = B$  as shown below.*

$$\begin{bmatrix} 2 & 3 & -4 \\ 3 & 4 & -5 \\ 5 & 0 & -6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}$$

## 2.2 Systems of Linear Equations; Gauss-Jordan Method

In this section you will learn to

1. Represent a system of linear equations as an augmented matrix
2. Solve the system using elementary row operations.

In this section, we learn to solve systems of linear equations using a process called the Gauss-Jordan method. The process begins by first expressing the system as a matrix, and then reducing it to an equivalent system by simple row operations. The process is continued until the solution is obvious from the matrix. The matrix that represents the system is called the augmented matrix, and the arithmetic manipulation that is used to move from a system to a reduced equivalent system is called a row operation.

**Example 2.2.1.** *Write the following system as an augmented matrix.*

$$\begin{aligned} 2x + 3y - 4z &= 5 \\ 3x + 4y - 5z &= -6 \\ 4x + 5y - 6z &= 7 \end{aligned}$$

**Solution 2.2.1.** *We express the above information in matrix form. Since a system is entirely determined by its coefficient matrix and by its matrix of constant terms, the augmented matrix will include only the coefficient matrix and the constant matrix. So the augmented matrix we get is as follows:*

$$\left[ \begin{array}{ccc|c} 2 & 3 & -4 & 5 \\ 3 & 4 & -5 & -6 \\ 4 & 5 & -6 & 7 \end{array} \right]$$



## 2.2. SYSTEMS OF LINEAR EQUATIONS; GAUSS-JORDAN METHOD41

In the last section, we expressed the system of equations as  $AX = B$ , where  $A$  represented the coefficient matrix, and  $B$  the matrix of constant terms. As an augmented matrix, we write the matrix as  $[A|B]$ . It is clear that all of the information is maintained in this matrix form, and only the letters  $x$ ,  $y$ , and  $z$  are missing. A student may choose to write  $x$ ,  $y$ , and  $z$  on top of the first three columns to help ease the transition.

**Example 2.2.2.** *For the following augmented matrix, write the system of equations it represents.*

$$\left[ \begin{array}{ccc|c} 1 & 3 & -5 & 2 \\ 2 & 0 & -3 & -5 \\ 3 & 2 & -3 & -1 \end{array} \right]$$

**Solution 2.2.2.** *The system is readily obtained as below.*

$$\begin{aligned} x + 3y - 5z &= 2 \\ 2x - 3z &= -5 \\ 3x + 2y - 3z &= -1 \end{aligned}$$

Once a system is expressed as an augmented matrix, the Gauss-Jordan method reduces the system into a series of equivalent systems by using the row operations. This row reduction continues until the system is expressed in what is called the reduced row echelon form. The reduced row echelon form of the coefficient matrix has 1's along the main diagonal and zeros elsewhere. The solution is readily obtained from this form.

The method is not much different from the algebraic operations we employed in the elimination method in the first chapter. The basic difference is that it is algorithmic in nature, and, therefore, can easily be programmed on a computer.

We will next solve a system of two equations with two unknowns, using the elimination method, and then show that the method is analogous to the Gauss-Jordan method.

**Example 2.2.3.** *Solve the following system by the elimination method.*

$$\begin{aligned} x + 3y &= 7 \\ 3x + 4y &= 11 \end{aligned}$$

**Solution 2.2.3.** We multiply the first equation by  $-3$  and add it to the second equation.

$$-3x - 9y = -21$$

$$3x + 4y = 11$$

This transforms our original system into an equivalent system:

$$x + 3y = 7$$

$$-5y = -10$$

Dividing the second equation by  $-5$ , we get the next equivalent system.

$$x + 3y = 7$$

$$y = 2$$

Multiplying the second equation by  $-3$  and adding it to the first, we get

$$x = 1$$

$$y = 2$$

**Example 2.2.4.** Solve the following system from Example 3 by the Gauss-Jordan method, and show the similarities in both methods by writing the equations next to the matrices.

$$x + 3y = 7$$

$$3x + 4y = 11$$

**Solution 2.2.4.** The augmented matrix for the system is as follows.

$$\left[ \begin{array}{cc|c} 1 & 3 & 7 \\ 3 & 4 & 11 \end{array} \right] \Leftrightarrow \begin{array}{l} x + 3y = 7 \\ 3x + 4y = 11 \end{array}$$

We multiply the first row by  $-3$  and add it to the second row.

$$\left[ \begin{array}{cc|c} 1 & 3 & 7 \\ 0 & -5 & -10 \end{array} \right] \Leftrightarrow \begin{array}{l} x + 3y = 7 \\ -5y = -10 \end{array}$$

Dividing the second row by  $-5$ , we get,

$$\left[ \begin{array}{cc|c} 1 & 3 & 7 \\ 0 & 1 & 2 \end{array} \right] \Leftrightarrow \begin{array}{l} x + 3y = 7 \\ y = 2 \end{array}$$

Finally, we multiply the second row by  $-3$  and add to the first row, and we get,

$$\left[ \begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array} \right] \Leftrightarrow \begin{array}{l} x = 1 \\ y = 2 \end{array}$$

### 2.2.1 Row Operations in Gauss-Jordan Method

The Gauss-Jordan method employs three fundamental row operations:

1. Any two rows in the augmented matrix may be interchanged.
2. Any row may be multiplied by a non-zero constant.
3. A constant multiple of a row may be added to another row.

One can easily see that these three row operations may make the system look different, but they do not change the solution of the system.

#### Example of Row Interchange

Consider the system of equations with two unknowns:

$$\begin{aligned}x + 3y &= 7 \\ 3x + 4y &= 11\end{aligned}$$

If we interchange the rows, we get:

$$\begin{aligned}3x + 4y &= 11 \\ x + 3y &= 7\end{aligned}$$

Clearly, this system has the same solution as the original.

#### Example of Multiplying a Row by a Constant

Consider the system again:

$$\begin{aligned}x + 3y &= 7 \\ 3x + 4y &= 11\end{aligned}$$

Multiplying the first row by  $-3$ , we get:

$$\begin{aligned}-3x - 9y &= -21 \\ 3x + 4y &= 11\end{aligned}$$

Once again, this new system has the same solution as the original.

**Example of Adding a Constant Multiple of One Row to Another**

For the system:

$$\begin{aligned}x + 3y &= 7 \\ 3x + 4y &= 11\end{aligned}$$

If we multiply the first row by  $-3$  and add it to the second row, we get:

$$\begin{aligned}x + 3y &= 7 \\ -5y &= -10\end{aligned}$$

The solution remains unchanged.

Now that we understand how the three row operations work, it is time to introduce the Gauss-Jordan method to solve systems of linear equations. As mentioned earlier, the Gauss-Jordan method starts out with an augmented matrix, and by a series of row operations ends up with a matrix that is in the reduced row echelon form. A matrix is in the reduced row echelon form if the first nonzero entry in each row is a 1, and the columns containing these 1's have all other entries as zeros. The reduced row echelon form also requires that the leading entry in each row be to the right of the leading entry in the row above it, and the rows containing all zeros be moved down to the bottom. We state the Gauss-Jordan method as follows.

**Gauss-Jordan Method Steps**

Here are the steps of the Gauss-Jordan method for solving linear systems:

1. Write the augmented matrix.
2. Interchange rows if necessary to obtain a non-zero number in the first row, first column.
3. Use a row operation to get a 1 as the entry in the first row and first column.
4. Use row operations to make all other entries as zeros in column one.
5. Interchange rows if necessary to obtain a nonzero number in the second row, second column. Use a row operation to make this entry 1. Use row operations to make all other entries as zeros in column two.

## 2.2. SYSTEMS OF LINEAR EQUATIONS; GAUSS-JORDAN METHOD 45

6. Repeat step 5 for row 3, column 3. Continue moving along the main diagonal until you reach the last row, or until the number is zero.
7. The final matrix is called the reduced row-echelon form.

**Example 2.2.5.** Solve the following system by the Gauss-Jordan method:

$$\begin{array}{rrcr} 2x & + & y & + & 2z & = & 10 \\ x & + & 2y & + & z & = & 8 \\ 3x & + & y & - & z & = & 2 \end{array}$$

**Solution 2.2.5.** We write the augmented matrix.

$$\left[ \begin{array}{ccc|c} 2 & 1 & 2 & 10 \\ 1 & 2 & 1 & 8 \\ 3 & 1 & -1 & 2 \end{array} \right]$$

We want a 1 in row one, column one. This can be obtained by dividing the first row by 2, or interchanging the second row with the first. Interchanging the rows is a better choice because that way we avoid fractions.

$$\left[ \begin{array}{ccc|c} 1 & 2 & 1 & 8 \\ 2 & 1 & 2 & 10 \\ 3 & 1 & -1 & 2 \end{array} \right] \text{ we interchanged row 1(R1) and row 2(R2)}$$

We need to make all other entries zeros in column 1. To make the entry (2) a zero in row 2, column 1, we multiply row 1 by -2 and add it to the second row. We get,

$$\left[ \begin{array}{ccc|c} 1 & 2 & 1 & 8 \\ 0 & -3 & 0 & -6 \\ 3 & 1 & -1 & 2 \end{array} \right] \quad -2R1 + R2$$

To make the entry (3) a zero in row 3, column 1, we multiply row 1 by -3 and add it to the third row. We get,

$$\left[ \begin{array}{ccc|c} 1 & 2 & 1 & 8 \\ 0 & -3 & 0 & -6 \\ 0 & -5 & -4 & -22 \end{array} \right] \quad -3R1 + R3$$

So far we have made a 1 in the left corner and all other entries zeros in that column. Now we move to the next diagonal entry, row 2, column 2. We need

to make this entry(-3) a 1 and make all other entries in this column zeros. To make row 2, column 2 entry a 1, we divide the entire second row by -3.

$$\left[ \begin{array}{ccc|c} 1 & 2 & 1 & 8 \\ 0 & 1 & 0 & 2 \\ 0 & -5 & -4 & -22 \end{array} \right] R2 \cdot \frac{1}{(-3)}$$

Next, we make all other entries zeros in the second column.

$$\left[ \begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & -4 & -12 \end{array} \right] -2R2 + R1 \text{ and } 5R2 + R3$$

We make the last diagonal entry a 1, by dividing row 3 by -4.

$$\left[ \begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array} \right] R3 \cdot \frac{1}{(-4)}$$

Finally, we make all other entries zeros in column 3.

$$\left[ \begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array} \right] -R3 + R1$$

Clearly, the solution reads  $x = 1$ ,  $y = 2$ , and  $z = 3$ .

Before we leave this section, we mention some terms we may need in the fourth chapter.

The process of obtaining a 1 in a location, and then making all other entries zeros in that column, is called pivoting.

The number that is made a 1 is called the pivot element, and the row that contains the pivot element is called the pivot row.

We often multiply the pivot row by a number and add it to another row to obtain a zero in the latter. The row to which a multiple of pivot row is added is called the target row.

## 2.3 Systems of Linear Equations – Special Cases

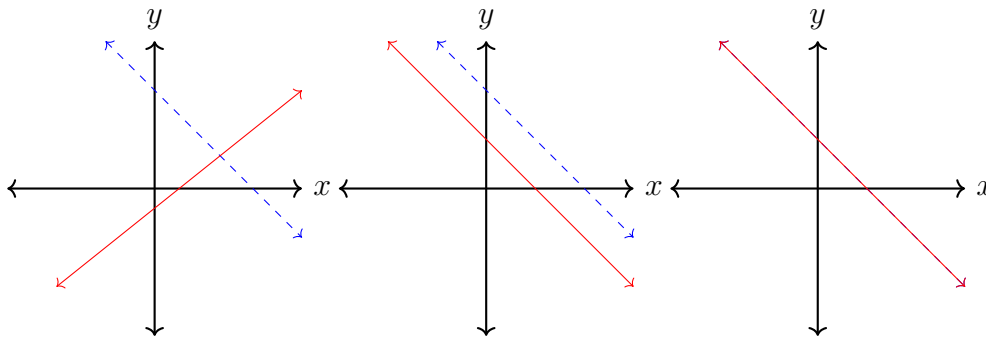
In this section you will learn to:

1. Determine the linear systems that have no solution.
2. Solve the linear systems that have infinitely many solutions.

If we consider the intersection of two lines in a plane, three things can happen.

1. The lines intersect in exactly one point. This is called an independent system.
2. The lines are parallel, so they do not intersect. This is called an inconsistent system.
3. The lines coincide; they intersect at infinitely many points. This is a dependent system.

The figures below show all three cases:



Every system of equations has either one solution, no solution, or infinitely many solutions.

In the last section, we used the Gauss-Jordan method to solve systems that had exactly one solution. In this section, we will determine the systems that have no solution, and solve the systems that have infinitely many solutions.

**Example 2.3.1.** *Solve the following system of equations using the Gauss-Jordan method:*

$$x + y = 7$$

$$x + y = 9$$

**Solution 2.3.1.** *Let us use the Gauss-Jordan method to solve this system. The augmented matrix is*

$$\left[ \begin{array}{cc|c} 1 & 1 & 7 \\ 1 & 1 & 9 \end{array} \right]$$

*If we multiply the first row by  $-1$  and add it to the second row, we get*

$$\left[ \begin{array}{cc|c} 1 & 1 & 7 \\ 0 & 0 & 2 \end{array} \right]$$

*Since 0 cannot equal 2, the last equation cannot be true for any choices of  $x$  and  $y$ . Alternatively, it is clear that the two lines are parallel; therefore, they do not intersect.*

In the examples that follow, we are going to start using a calculator to row reduce the augmented matrix, in order to focus on understanding the answer rather than focusing on the process of carrying out the row operations.

**Example 2.3.2.** *Solve the following system of equations:*

$$\begin{aligned} 2x + 3y - 4z &= 7 \\ 3x + 4y - 2z &= 9 \\ 5x + 7y - 6z &= 20 \end{aligned}$$

**Solution 2.3.2.** *We represent the system as an augmented matrix:*

$$\left[ \begin{array}{ccc|c} 2 & 3 & -4 & 7 \\ 3 & 4 & -2 & 9 \\ 5 & 7 & -6 & 20 \end{array} \right]$$

*By obtaining the reduced row-echelon form from a matrix calculator, we get:*

$$\left[ \begin{array}{ccc|c} 1 & 0 & 10 & 0 \\ 0 & 1 & -8 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right]$$

*The bottom row implies  $0x + 0y + 0z = 1$ , which is a contradiction. Thus, the system is inconsistent and has no solution.*



**Example 2.3.3.** Solve the following system of equations:

$$x + y = 7$$

$$x + y = 7$$

**Solution 2.3.3.** The problem asks for the intersection of two identical lines, meaning the lines coincide and intersect at an infinite number of points.

A few intersection points are listed as follows:  $(3, 4)$ ,  $(5, 2)$ ,  $(-1, 8)$ ,  $(-6, 13)$ , etc. However, when a system has an infinite number of solutions, the solution is often expressed in parametric form. This can be done by assigning an arbitrary constant,  $t$ , to one of the variables and solving for the remaining variables. If we let  $y = t$ , then  $x = 7 - t$ . In other words, all ordered pairs of the form  $(7 - t, t)$  satisfy the given system of equations.

Alternatively, solving with the Gauss-Jordan method, we obtain the reduced row-echelon form below, which includes a row of all zeros that can be ignored since it provides no additional information about the values of  $x$  and  $y$  that solve the system.

$$\left[ \begin{array}{cc|c} 1 & 1 & 7 \\ 0 & 0 & 0 \end{array} \right]$$

This leaves us with only one equation but two variables. Whenever there are more variables than equations, the solution must be expressed as a parametric solution in terms of an arbitrary constant, as shown above.

Parametric Solution:  $x = 7 - t$ ,  $y = t$ .

**Example 2.3.4.** Solve the following system of equations:

$$x + y + z = 2$$

$$2x + y - z = 3$$

$$3x + 2y = 5$$

**Solution 2.3.4.** The augmented matrix and the reduced row-echelon form are given below:

$$\left[ \begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 2 & 1 & -1 & 3 \\ 3 & 2 & 0 & 5 \end{array} \right] \rightarrow \left[ \begin{array}{ccc|c} 1 & 0 & -2 & 1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

Since the last equation dropped out, we are left with two equations and three variables. This means the system has an infinite number of solutions. We express those solutions in the parametric form by letting the last variable  $z$  equal the parameter  $t$ .

The first equation reads  $x - 2z = 1$ , therefore,  $x = 1 + 2z$ . The second equation reads  $y + 3z = 1$ , therefore,  $y = 1 - 3z$ . And now if we let  $z = t$ , the parametric solution is expressed as follows:

$$\text{Parametric Solution: } x = 1 + 2t, \quad y = 1 - 3t, \quad z = t.$$

The reader should note that particular solutions, or specific solutions, to the system can be obtained by assigning values to the parameter  $t$ . For example:

- If we let  $t = 2$ , we have the solution  $x = 5, y = -5, z = 2$ :  $(5, -5, 2)$ .
- If we let  $t = 0$ , we have the solution  $x = 1, y = 1, z = 0$ :  $(1, 1, 0)$ .

**Example 2.3.5.** Solve the following system of equations:

$$\begin{aligned} x + 2y - 3z &= 5 \\ 2x + 4y - 6z &= 10 \\ 3x + 6y - 9z &= 15 \end{aligned}$$

**Solution 2.3.5.** The reduced row-echelon form is given below:

$$\left[ \begin{array}{ccc|c} 1 & 2 & -3 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

This time the last two equations drop out. We are left with one equation and three variables. Again, there are an infinite number of solutions. But this time the answer must be expressed in terms of two arbitrary constants.

If we let  $z = t$  and  $y = s$ , the first equation  $x + 2y - 3z = 5$  results in  $x = 5 - 2s + 3t$ . We rewrite the parametric solution as:

$$\text{Parametric Solution: } x = 5 - 2s + 3t, \quad y = s, \quad z = t.$$

**Summary 2.3.1: Systems of Equations - Special Cases**

1. If any row of the reduced row-echelon form of the matrix gives a false statement such as  $0 = 1$ , the system is inconsistent and has no solution.
2. If the reduced row echelon form has fewer equations than the variables and the system is consistent, then the system has an infinite number of solutions. Remember the rows that contain all zeros are dropped.
  - (a) If a system has an infinite number of solutions, the solution must be expressed in the parametric form.
  - (b) The number of arbitrary parameters equals the number of variables minus the number of equations.

## 2.4 Inverse Matrices

In this section you will learn to:

1. Find the inverse of a matrix, if it exists.
2. Use inverses to solve linear systems.

In this section, we will learn to find the inverse of a matrix, if it exists. Later, we will use matrix inverses to solve linear systems.

**Definition 2.4.1.** An  $n \times n$  matrix has an **inverse** if there exists a matrix  $B$  such that  $AB = BA = I_n$ , where  $I_n$  is an  $n \times n$  identity matrix. The **inverse** of a matrix  $A$ , if it exists, is denoted by the symbol  $A^{-1}$ .

**Example 2.4.1.** Given matrices  $A$  and  $B$  below, verify that they are inverses.

$$A = \begin{bmatrix} 4 & 1 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ -3 & 4 \end{bmatrix}$$

**Solution 2.4.1.** The matrices are inverses if the product  $AB$  and  $BA$  both equal the identity matrix of dimension  $2 \times 2$ , denoted as  $I_2$ :

$$AB = \begin{bmatrix} 4 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

and

$$BA = \begin{bmatrix} 1 & -1 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

Clearly, that is the case; therefore, the matrices  $A$  and  $B$  are inverses of each other.

**Example 2.4.2.** Find the inverse of the matrix  $A = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}$ .

**Solution 2.4.2.** Suppose  $A$  has an inverse, and it is denoted as  $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ . Then  $AB = I_2$ :

$$\begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

After multiplying the matrices on the left side, we get the system:

$$\begin{aligned} 3a + c &= 1 \\ 3b + d &= 0 \\ 5a + 2c &= 0 \\ 5b + 2d &= 1 \end{aligned}$$

Solving this system, we find  $a = 2$ ,  $b = -1$ ,  $c = -5$ , and  $d = 3$ . Therefore, the inverse of matrix  $A$  is  $B = \begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix}$ .

In this problem, finding the inverse of matrix  $A$  amounted to solving the system of equations:

$$\begin{aligned} 3a + c &= 1 \\ 3b + d &= 0 \\ 5a + 2c &= 0 \\ 5b + 2d &= 1 \end{aligned}$$

Actually, it can be written as two systems, one with variables  $a$  and  $c$ , and the other with  $b$  and  $d$ . The augmented matrices for both are given below.

$$\left[ \begin{array}{cc|c} 3 & 1 & 1 \\ 5 & 2 & 0 \end{array} \right] \quad \text{and} \quad \left[ \begin{array}{cc|c} 3 & 1 & 0 \\ 5 & 2 & 1 \end{array} \right]$$

As we look at the two augmented matrices, we notice that the coefficient matrix for both the matrices is the same. This implies the row operations of

the Gauss-Jordan method will also be the same. A great deal of work can be saved if the two right-hand columns are grouped together to form one augmented matrix as below.

$$\left[ \begin{array}{cc|cc} 3 & 1 & 1 & 0 \\ 5 & 2 & 0 & 1 \end{array} \right]$$

And solving this system, we get

$$\left[ \begin{array}{cc|cc} 1 & 0 & 2 & -1 \\ 0 & 1 & -5 & 3 \end{array} \right]$$

The matrix on the right side of the vertical line is the  $A^{-1}$  matrix. What you just witnessed is no coincidence. This is the method that is often employed in finding the inverse of a matrix.

#### Summary 2.4.1:

The Method for Finding the Inverse of a Matrix

1. Write the augmented matrix  $[A|I_n]$ .
2. Write the augmented matrix in step 1 in reduced row echelon form.
3. If the reduced row echelon form in 2 is  $[I_n|B]$ , then  $B$  is the inverse of  $A$ .
4. If the left side of the row reduced echelon is not an identity matrix, the inverse does not exist.

**Example 2.4.3.** *Given the matrix  $A$  below, find its inverse.*

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 3 & 0 \\ 0 & -2 & 1 \end{bmatrix}$$

**Solution 2.4.3.** *We write the augmented matrix as follows.*

$$\left[ \begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 2 & 3 & 0 & 0 & 1 & 0 \\ 0 & -2 & 1 & 0 & 0 & 1 \end{array} \right]$$

We will reduce this matrix using the Gauss-Jordan method. Multiplying the first row by  $-2$  and adding it to the second row, we get

$$\left[ \begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 5 & -2 & -2 & 1 & 0 \\ 0 & -2 & 1 & 0 & 0 & 1 \end{array} \right]$$

If we swap the second and third rows, we get

$$\left[ \begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 & 1 \\ 0 & 5 & -2 & -2 & 1 & 0 \end{array} \right]$$

Divide the second row by  $-2$ . The result is

$$\left[ \begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1/2 & 0 & 0 & -1/2 \\ 0 & 5 & -2 & -2 & 1 & 0 \end{array} \right]$$

Let us do two operations here. 1) Add the second row to the first. 2) Add  $-5$  times the second row to the third. And we get

$$\left[ \begin{array}{ccc|ccc} 1 & 0 & 1/2 & 1 & 0 & -1/2 \\ 0 & 1 & -1/2 & 0 & 0 & -1/2 \\ 0 & 0 & 1/2 & -2 & 1 & 5/2 \end{array} \right]$$

Multiplying the third row by 2 results in

$$\left[ \begin{array}{ccc|ccc} 1 & 0 & 1/2 & 1 & 0 & -1/2 \\ 0 & 1 & -1/2 & 0 & 0 & -1/2 \\ 0 & 0 & 1 & -4 & 2 & 5 \end{array} \right]$$

Multiply the third row by  $1/2$  and add it to the second. Also, multiply the third row by  $-1/2$  and add it to the first.

$$\left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -1 & -3 \\ 0 & 1 & 0 & -2 & 1 & 2 \\ 0 & 0 & 1 & -4 & 2 & 5 \end{array} \right]$$

Therefore, the inverse of matrix  $A$  is  $A^{-1} = \begin{bmatrix} 3 & -1 & -3 \\ -2 & 1 & 2 \\ -4 & 2 & 5 \end{bmatrix}$ .

One should verify the result by multiplying the two matrices to see if the product does, indeed, equal the identity matrix.

Now that we know how to find the inverse of a matrix, we will use inverses to solve systems of equations. The method is analogous to solving a simple equation like the one below.

$$\frac{2}{3}x = 4$$

**Example 2.4.4.** Solve the following equation:

$$x = 4$$

**Solution 2.4.4.** To solve the above equation, we multiply both sides of the equation by the multiplicative inverse of  $\frac{2}{3}$ , which happens to be  $\frac{3}{2}$ . We get

$$\frac{3}{2} \cdot \frac{2}{3}x = 4 \cdot \frac{3}{2}$$

Hence,

$$x = 6.$$

We use example 2.4.4 as an analogy to show how linear systems of the form  $AX = B$  are solved. To solve a linear system, we first write the system in the matrix equation  $AX = B$ , where  $A$  is the coefficient matrix,  $X$  is the matrix of variables, and  $B$  is the matrix of constant terms. We then multiply both sides of this equation by the multiplicative inverse of the matrix  $A$ . Consider the following example.

**Example 2.4.5.** Solve the following system

$$\begin{aligned} 3x + y &= 3 \\ 5x + 2y &= 4 \end{aligned}$$

**Solution 2.4.5.** To solve the above equation, first we express the system as

$$AX = B$$

where  $A$  is the coefficient matrix, and  $B$  is the matrix of constant terms. We get

$$\begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

To solve this system, we multiply both sides of the matrix equation  $AX = B$  by  $A^{-1}$ . Matrix multiplication is not commutative, so we need to multiply by  $A^{-1}$  on the left on both sides of the equation.

Matrix  $A$  is the same matrix  $A$  whose inverse we found in Example 2.4.2, so  $A^{-1} = \begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix}$ .

Multiplying both sides by  $A^{-1}$ , we get

$$\begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$$

Therefore,  $x = 2$ , and  $y = -3$ .

**Example 2.4.6.** Solve the following system:

$$\begin{aligned} x - y + z &= 6 \\ 2x + 3y &= 1 \\ -2y + z &= 5 \end{aligned}$$

**Solution 2.4.6.** To solve the above equation, we write the system in matrix form  $AX = B$  as follows:

$$\begin{bmatrix} 1 & -1 & 1 \\ 2 & 3 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \\ 5 \end{bmatrix}$$



To solve this system, we need the inverse of  $A$ . From Example 2.4.3,  $A^{-1}$  is given by

$$A^{-1} = \begin{bmatrix} 3 & -1 & -3 \\ -2 & 1 & 2 \\ -4 & 2 & 5 \end{bmatrix}$$

Multiplying both sides of the matrix equation  $AX = B$  on the left by  $A^{-1}$ , we get

$$\begin{bmatrix} 3 & -1 & -3 \\ -2 & 1 & 2 \\ -4 & 2 & 5 \end{bmatrix} \begin{bmatrix} 6 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

After multiplying the matrices, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$$

Therefore,  $x = 2$ ,  $y = -1$ , and  $z = 3$ .

We remind the reader that not every system of equations can be solved by the matrix inverse method. Although the Gauss-Jordan method works for every situation, the matrix inverse method works only in cases where the inverse of the square matrix exists. In such cases the system has a unique solution.

#### Summary 2.4.2: Finding the Inverse of a Matrix

1. Write the augmented matrix  $[A|I_n]$ .
2. Write the augmented matrix in step 1 in reduced row echelon form.
3. If the reduced row echelon form in step 2 is  $[I_n|B]$ , then  $B$  is the inverse of  $A$ .
4. If the left side of the row reduced echelon is not an identity matrix, the inverse does not exist.

**Summary 2.4.3: Solving a System of Equations When a Unique Solution Exists**

1. Express the system in the matrix equation  $AX = B$ .
2. To solve the equation  $AX = B$ , multiply both sides by  $A^{-1}$ :

$$AX = B$$

$$A^{-1}AX = A^{-1}B$$

$$IX = A^{-1}B \quad \text{where } I \text{ is the identity matrix}$$

## 2.5 Application of Matrices in Cryptography

In this section, you will learn to:

1. Encode a message using matrix multiplication.
2. Decode a coded message using the matrix inverse and matrix multiplication.

Encryption dates back approximately 4000 years. Historical accounts indicate that the Chinese, Egyptians, Indians, and Greeks encrypted messages in some way for various purposes. One famous encryption scheme is called the Caesar cipher, also called a substitution cipher, used by Julius Caesar, involved shifting letters in the alphabet, such as replacing A by C, B by D, C by E, etc., to encode a message. Substitution ciphers are too simple in design to be considered secure today.

In the middle ages, European nations began to use encryption. A variety of encryption methods were used in the US from the Revolutionary War, through the Civil War, and on into modern times.

Applications of mathematical theory and methods to encryption became widespread in military usage in the 20th century. The military would encode messages before sending, and the recipient would decode the message, in order to send information about military operations in a manner that kept the information safe if the message was intercepted. In World War II, encryption played an important role, as both Allied and Axis powers sent encrypted messages and devoted significant resources to strengthening their own encryption while also trying to break the opposition's encryption.

In this section, we will examine a method of encryption that uses matrix multiplication and matrix inverses. This method, known as the Hill Algorithm, was created by Lester Hill, a mathematics professor who taught at several US colleges and also was involved with military encryption. The Hill algorithm marks the introduction of modern mathematical theory and methods to the field of cryptography.

These days, the Hill Algorithm is not considered a secure encryption method; it is relatively easy to break with modern technology. However, in 1929 when it was developed, modern computing technology did not exist. This method, which we can handle easily with today's technology, was too cumbersome to use with hand calculations. Hill devised a mechanical encryption machine to help with the mathematics; his machine relied on gears and levers but never gained widespread use. Hill's method was considered sophisticated and powerful in its time and is one of many methods influencing techniques in use today. Other encryption methods at that time also utilized special coding machines. Alan Turing, a computer scientist pioneer in the field of artificial intelligence, invented a machine that was able to decrypt messages encrypted by the German Enigma machine, helping to turn the tide of World War II.

With the advent of the computer age and internet communication, the use of encryption has become widespread in communication and in keeping private data secure; it is no longer limited to military uses. Modern encryption methods are more complicated, often combining several steps or methods to encrypt data to keep it more secure and harder to break. Some modern methods make use of matrices as part of the encryption and decryption process; other fields of mathematics such as number theory play a large role in modern cryptography.

### 2.5.1 Using Matrices for Encoding and Decoding

To use matrices in encoding and decoding secret messages, our procedure is as follows.

We first convert the secret message into a string of numbers by arbitrarily assigning a number to each letter of the message. Next, we convert this string of numbers into a new set of numbers by multiplying the string by a square matrix of our choice that has an inverse. This new set of numbers represents

the coded message.

To decode the message, we take the string of coded numbers and multiply it by the inverse of the matrix to get the original string of numbers. Finally, by associating the numbers with their corresponding letters, we obtain the original message.

In this section, we will use the correspondence shown below where letters A to Z correspond to the numbers 1 to 26, a space is represented by the number 27, and punctuation is ignored.

A	B	C	D	E	F	G	H	I	J	K	L	M
1	2	3	4	5	6	7	8	9	10	11	12	13

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
14	15	16	17	18	19	20	21	22	23	24	25	26

**Example 2.5.1.** Use matrix  $A$  to encode the message: *ATTACK NOW!*

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$$

**Solution 2.5.1.** We divide the letters of the message into groups of two.

$$AT \quad TA \quad CK \quad \_N \quad OW$$

We assign the numbers to these letters from the above table, and convert each pair of numbers into  $2 \times 1$  matrices. In the case where a single letter is left over on the end, a space is added to make it into a pair.

$$\begin{bmatrix} A \\ T \end{bmatrix} = \begin{bmatrix} 1 \\ 20 \end{bmatrix}, \quad \begin{bmatrix} T \\ A \end{bmatrix} = \begin{bmatrix} 20 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} C \\ K \end{bmatrix} = \begin{bmatrix} 3 \\ 11 \end{bmatrix}, \quad \begin{bmatrix} \_ \\ N \end{bmatrix} = \begin{bmatrix} 27 \\ 14 \end{bmatrix}, \quad \begin{bmatrix} O \\ W \end{bmatrix} = \begin{bmatrix} 15 \\ 23 \end{bmatrix}$$

So at this stage, our message expressed as  $2 \times 1$  matrices is as follows.

$$\begin{bmatrix} 1 \\ 20 \end{bmatrix}, \quad \begin{bmatrix} 20 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 3 \\ 11 \end{bmatrix}, \quad \begin{bmatrix} 27 \\ 14 \end{bmatrix}, \quad \begin{bmatrix} 15 \\ 23 \end{bmatrix}$$

Now to encode, we multiply, on the left, each matrix of our message by the matrix  $A$ .

For example, the product of  $A$  with our first matrix is:

$$A \cdot \begin{bmatrix} 1 \\ 20 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 20 \end{bmatrix} = \begin{bmatrix} 41 \\ 61 \end{bmatrix}$$

And the product of  $A$  with our second matrix is:

$$A \cdot \begin{bmatrix} 20 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 20 \\ 1 \end{bmatrix} = \begin{bmatrix} 22 \\ 23 \end{bmatrix}$$

Multiplying matrix  $A$  by each matrix in our list, in turn, gives the desired coded message:

$$\begin{bmatrix} 41 \\ 61 \end{bmatrix}, \quad \begin{bmatrix} 22 \\ 23 \end{bmatrix}, \quad \begin{bmatrix} 25 \\ 36 \end{bmatrix}, \quad \begin{bmatrix} 55 \\ 69 \end{bmatrix}, \quad \begin{bmatrix} 61 \\ 84 \end{bmatrix}$$

**Example 2.5.2.** Decode the following message that was encoded using matrix

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}.$$

$$\begin{bmatrix} 21 \\ 26 \end{bmatrix}, \quad \begin{bmatrix} 37 \\ 53 \end{bmatrix}, \quad \begin{bmatrix} 45 \\ 54 \end{bmatrix}, \quad \begin{bmatrix} 74 \\ 101 \end{bmatrix}, \quad \begin{bmatrix} 53 \\ 69 \end{bmatrix}$$

**Solution 2.5.2.** Since this message was encoded by multiplying by the matrix  $A$  in Example 2.4.2, we decode this message by first multiplying each matrix, on the left, by the inverse of matrix  $A$  given below.

$$A^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$$

For example:

$$\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 21 \\ 26 \end{bmatrix} = \begin{bmatrix} 11 \\ 5 \end{bmatrix}$$

By multiplying each of the matrices in our list by the matrix  $A^{-1}$ , we get the following.

$$\begin{bmatrix} 11 \\ 5 \end{bmatrix}, \quad \begin{bmatrix} 5 \\ 16 \end{bmatrix}, \quad \begin{bmatrix} 27 \\ 9 \end{bmatrix}, \quad \begin{bmatrix} 20 \\ 27 \end{bmatrix}, \quad \begin{bmatrix} 21 \\ 16 \end{bmatrix}$$

Finally, by associating the numbers with their corresponding letters, we obtain:

$$\begin{bmatrix} K \\ E \end{bmatrix}, \quad \begin{bmatrix} E \\ P \end{bmatrix}, \quad \begin{bmatrix} - \\ T \end{bmatrix}, \quad \begin{bmatrix} I \\ U \end{bmatrix}, \quad \begin{bmatrix} P \\ - \end{bmatrix}$$

And the message reads: *KEEP IT UP*.

Now suppose we wanted to use a  $3 \times 3$  matrix to encode a message, then instead of dividing the letters into groups of two, we would divide them into groups of three.

**Example 2.5.3.** Using the matrix  $B = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \end{bmatrix}$ , encode the message:

*ATTACK NOW!*

**Solution 2.5.3.** We divide the letters of the message into groups of three.

*ATT ACK \_NO W\_*

Note that since the single letter "W" was left over on the end, we added two spaces to make it into a triplet.

Now we assign the numbers their corresponding letters from the table, and convert each triplet of numbers into  $3 \times 1$  matrices. We get

$$\begin{bmatrix} A \\ T \\ T \end{bmatrix} = \begin{bmatrix} 1 \\ 20 \\ 20 \end{bmatrix}, \quad \begin{bmatrix} A \\ C \\ K \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 11 \end{bmatrix}, \quad \begin{bmatrix} - \\ N \\ O \end{bmatrix} = \begin{bmatrix} 27 \\ 14 \\ 15 \end{bmatrix}, \quad \begin{bmatrix} W \\ - \\ - \end{bmatrix} = \begin{bmatrix} 23 \\ 27 \\ 27 \end{bmatrix}$$

So far we have,

$$\begin{bmatrix} 1 \\ 20 \\ 20 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 3 \\ 11 \end{bmatrix}, \quad \begin{bmatrix} 27 \\ 14 \\ 15 \end{bmatrix}, \quad \begin{bmatrix} 23 \\ 27 \\ 27 \end{bmatrix}$$

We multiply, on the left, each matrix of our message by the matrix  $B$ . For example,

$$\begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 20 \\ 20 \end{bmatrix} = \begin{bmatrix} 1 \\ 21 \\ 42 \end{bmatrix}$$

By multiplying each of the matrices in (III) by the matrix  $B$ , we get the desired coded message as follows:

$$\begin{bmatrix} 1 \\ 21 \\ 42 \end{bmatrix}, \quad \begin{bmatrix} -7 \\ 12 \\ 16 \end{bmatrix}, \quad \begin{bmatrix} 26 \\ 42 \\ 83 \end{bmatrix}, \quad \begin{bmatrix} 23 \\ 50 \\ 100 \end{bmatrix}$$

If we need to decode this message, we simply multiply the coded message by  $B^{-1}$ , and associate the numbers with the corresponding letters of the alphabet.

In Example 2.5.4 we will demonstrate how to use matrix  $B^{-1}$  to decode an encrypted message.

**Example 2.5.4.** Decode the following message that was encoded using matrix

$$B = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \end{bmatrix} :$$

$$\begin{bmatrix} 11 \\ 20 \\ 43 \end{bmatrix}, \quad \begin{bmatrix} 25 \\ 10 \\ 41 \end{bmatrix}, \quad \begin{bmatrix} 22 \\ 14 \\ 41 \end{bmatrix}$$

**Solution 2.5.4.** Since this message was encoded by multiplying by the matrix  $B$ . We first determine the inverse of  $B$ .

$$B^{-1} = \begin{bmatrix} 1 & 2 & -1 \\ -1 & -3 & 2 \\ -1 & -1 & 1 \end{bmatrix}$$

To decode the message, we multiply each matrix, on the left, by  $B^{-1}$ . For example,

$$\begin{bmatrix} 1 & 2 & -1 \\ -1 & -3 & 2 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 11 \\ 20 \\ 43 \end{bmatrix} = \begin{bmatrix} 8 \\ 15 \\ 12 \end{bmatrix}$$

Multiplying each of the matrices in our list by the matrix  $B^{-1}$  gives the following:

$$\begin{bmatrix} 8 \\ 15 \\ 12 \end{bmatrix}, \quad \begin{bmatrix} 4 \\ 27 \\ 6 \end{bmatrix}, \quad \begin{bmatrix} 9 \\ 18 \\ 5 \end{bmatrix}$$

Finally, by associating the numbers with their corresponding letters, we obtain:

$$\begin{bmatrix} H \\ O \\ L \end{bmatrix}, \quad \begin{bmatrix} D \\ - \\ F \end{bmatrix}, \quad \begin{bmatrix} I \\ R \\ E \end{bmatrix}$$

The message reads: *HOLD FIRE*.

#### Summary 2.5.1: Encoding and Decoding

##### To Encode a Message

1. Divide the letters of the message into groups of two or three.
2. Convert each group into a string of numbers by assigning a number to each letter of the message. Remember to assign letters to blank spaces.
3. Convert each group of numbers into column matrices.
4. Convert these column matrices into a new set of column matrices by multiplying them with a compatible square matrix of your choice that has an inverse. This new set of numbers or matrices represents the coded message.

##### To Decode a Message

1. Take the string of coded numbers and multiply it by the inverse of the matrix that was used to encode the message.
2. Associate the numbers with their corresponding letters.



## 2.6 Applications – Leontief Models

In this section you will learn

1. Application of matrices to model closed economic systems
2. Application of matrices to model open economic systems

In the 1930s, Wassily Wassilyevich Leontief (holder of one of the greatest names ever) used matrices to model economic systems. His models, often referred to as the input-output models, divide the economy into sectors where each sector produces goods and services not only for itself but also for other sectors. These sectors are dependent on each other, and the total input always equals the total output. In 1973, he won the Nobel Prize in Economics for his work in this field. In this section, we look at both the closed and the open models that he developed.

### 2.6.1 The Closed Model

As an example of the closed model, we look at a very simple economy, where there are only three sectors: food, shelter, and clothing.

**Example 2.6.1.** *We assume that in a village there is a farmer, carpenter, and a tailor, who provide the three essential goods: food, shelter, and clothing. Suppose the farmer himself consumes 40% of the food he produces, and gives 40% to the carpenter, and 20% to the tailor. Thirty percent of the carpenter's production is consumed by himself, 40% by the farmer, and 30% by the carpenter. Fifty percent of the tailor's production is used by himself, 30% by the farmer, and 20% by the tailor. Write the matrix that describes this closed model.*

**Solution 2.6.1.** *The table below describes the above information.*

	<i>Proportion produced by the farmer</i>	<i>Proportion produced by the carpenter</i>	<i>Proportion produced by the tailor</i>
<i>The proportion used by the farmer</i>	.40	.40	.30
<i>The proportion used by the carpenter</i>	.40	.30	.20
<i>The proportion used by the tailor</i>	.20	.30	.50

*In matrix form, it can be written as follows.*

$$A = \begin{bmatrix} .40 & .40 & .30 \\ .40 & .30 & .20 \\ .20 & .30 & .50 \end{bmatrix}$$

*This matrix is called the input-output matrix. It is important that we read the matrix correctly. For example, the entry  $A_{23}$ , the entry in row 2 and column 3, represents the following.*

$A_{23} = 20\%$  of the tailor's production is used by the carpenter.

$A_{33} = 50\%$  of the tailor's production is used by the tailor.

**Example 2.6.2.** *In Example 2.6.1 above, how much should each person get for his efforts?*

**Solution 2.6.2.** *We choose the following variables.*

$x = \text{Farmer's pay}$

$y = \text{Carpenter's pay}$

$z = \text{Tailor's pay}$

*As we said earlier, in this model input must equal output. That is, the amount paid by each equals the amount received by each.*

*Let us say the farmer gets paid  $x$  dollars. Let us now look at the farmer's expenses. The farmer uses up 40% of his own production, that is, of the  $x$  dollars he gets paid, he pays himself  $.40x$  dollars, he pays  $.40y$  dollars to the*

carpenter, and  $.30z$  to the tailor. Since the expenses equal the wages, we get the following equation.

$$x = .40x + .40y + .30z$$

In the same manner, we get

$$y = .40x + .30y + .20z$$

$$z = .20x + .30y + .50z$$

The above system can be written as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} .40 & .40 & .30 \\ .40 & .30 & .20 \\ .20 & .30 & .50 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

This system is often referred to as  $X = AX$ .

Simplification results in the system of equations  $(I - A)X = 0$

$$\begin{bmatrix} .60 & -.40 & -.30 \\ -.40 & .70 & -.20 \\ -.20 & -.30 & .50 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

We put this into an augmented matrix

$$\left[ \begin{array}{ccc|c} .60 & -.40 & -.30 & 0 \\ -.40 & .70 & -.20 & 0 \\ -.20 & -.30 & .50 & 0 \end{array} \right]$$

Solving for  $x, y$ , and  $z$  using the Gauss-Jordan method, we get

$$\left[ \begin{array}{ccc|c} 1 & 0 & -\frac{29}{26} & 0 \\ 0 & 1 & -\frac{12}{13} & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

This gives parametric equations:

$$x = \frac{29}{26}t, \quad y = \frac{12}{13}t, \quad z = t$$

Since we are only trying to determine the proportions of the pay, we can choose  $t$  to be any value. Suppose we let  $t = \$2600$ , then we get

$$x = \$2900, \quad y = \$2400, \quad z = \$2600$$

**Note 2.6.1.** *The use of a graphing calculator or computer application in solving the systems of linear matrix equations in these problems is strongly recommended.*

## 2.6.2 The Open Model

The open model is more realistic as it deals with the economy where sectors of the economy not only satisfy each other's needs but also satisfy some outside demands. In this case, the outside demands are put on by the consumer. But the basic assumption is still the same: whatever is produced is consumed.

Let us again look at a very simple scenario. Suppose the economy consists of three people: the farmer F, the carpenter C, and the tailor T. A part of the farmer's production is used by all three, and the rest is used by the consumer. In the same manner, a part of the carpenter's and the tailor's production is used by all three, and the rest is used by the consumer.

Let us assume that whatever the farmer produces, 20% is used by him, 15% by the carpenter, 10% by the tailor, and the consumer uses the other \$40 billion worth of food. Ten percent of the carpenter's production is used by him, 25% by the farmer, 5% by the tailor, and \$50 billion worth by the consumer. Fifteen percent of the clothing is used by the tailor, 10% by the farmer, 5% by the carpenter, and the remaining \$60 billion worth by the consumer. We write the internal consumption in the following table and express the demand as the matrix D.

	F produces	C produces	T produces
F uses	0.20	0.25	0.10
C uses	0.15	0.10	0.05
T uses	0.10	0.05	0.15

The consumer demand for each industry in billions of dollars is given by the matrix  $D = \begin{bmatrix} 40 \\ 50 \\ 60 \end{bmatrix}$ .

**Example 2.6.3.** *In the example above, what should be, in billions of dollars, the required output by each industry to meet the demand given by the matrix D?*

**Solution 2.6.3.** We choose the following variables.

$$\begin{aligned}x &= \text{Farmer's output} \\y &= \text{Carpenter's output} \\z &= \text{Tailor's output}\end{aligned}$$

In the closed model, our equation was  $X = AX$ , that is, the total input equals the total output. This time our equation is similar with the exception of the demand by the consumer.

So our equation for the open model should be  $X = AX + D$ , where  $D$  represents the demand matrix.

We express it as follows:

$$\begin{aligned}X &= AX + D \\ \begin{bmatrix} x \\ y \\ z \end{bmatrix} &= \begin{bmatrix} .20 & .25 & .10 \\ .15 & .10 & .05 \\ .10 & .05 & .15 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 40 \\ 50 \\ 60 \end{bmatrix}\end{aligned}$$

To solve this system, we write it as

$$\begin{aligned}X &= AX + D \\ (I - A)X &= D \\ X &= (I - A)^{-1}D\end{aligned}$$

where  $I$  is a  $3 \times 3$  identity matrix.

$$\begin{aligned}I - A &= \begin{bmatrix} .80 & -.25 & -.10 \\ -.15 & .90 & -.05 \\ -.10 & -.05 & .85 \end{bmatrix} \\ (I - A)^{-1} &= \begin{bmatrix} 1.3445 & .3835 & .1807 \\ .2336 & 1.1814 & .097 \\ .1719 & .1146 & 1.2034 \end{bmatrix} \\ X &= \begin{bmatrix} 1.3445 & .3835 & .1807 \\ .2336 & 1.1814 & .097 \\ .1719 & .1146 & 1.2034 \end{bmatrix} \begin{bmatrix} 40 \\ 50 \\ 60 \end{bmatrix}\end{aligned}$$

$$X = \begin{bmatrix} 83.7999 \\ 74.2341 \\ 84.8138 \end{bmatrix}$$

The three industries must produce the following amount of goods in billions of dollars.

$$\text{Farmer} = 83.7999$$

$$\text{Carpenter} = 74.2341$$

$$\text{Tailor} = 84.8138$$

We will do one more problem like the one above, except this time we give the amount of internal and external consumption in dollars and ask for the proportion of the amounts consumed by each of the industries. In other words, we ask for the matrix  $A$ .

**Example 2.6.4.** Suppose an economy consists of three industries  $F$ ,  $C$ , and  $T$ . Each of the industries produces for internal consumption among themselves, as well as for external demand by the consumer. The table shows the use of each industry's production in dollars.

	<b><math>F</math></b>	<b><math>C</math></b>	<b><math>T</math></b>	<b><math>Demand</math></b>	<b><math>Total</math></b>
<b><math>F</math></b>	40	50	60	100	250
<b><math>C</math></b>	30	40	40	110	220
<b><math>T</math></b>	20	30	30	120	200

The first row says that of the \$250 dollars worth of production by the industry  $F$ , \$40 is used by  $F$ , \$50 is used by  $C$ , \$60 is used by  $T$ , and the remainder of \$100 is used by the consumer. The other rows are described in a similar manner.

Once again, the total input equals the total output. Find the proportion of the amounts consumed by each of the industries. In other words, find the matrix  $A$ .

**Solution 2.6.4.** We are being asked to determine the following:

How much of the production of each of the three industries,  $F$ ,  $C$ , and  $T$  is required to produce one unit of  $F$ ? The same way, how much of the production

of each of the three industries,  $F$ ,  $C$ , and  $T$  is required to produce one unit of  $C$ ? And finally, how much of the production of each of the three industries,  $F$ ,  $C$ , and  $T$  is required to produce one unit of  $T$ ?

Since we are looking for proportions, we need to divide the production of each industry by the total production for each industry.

We analyze as follows: To produce 250 units of  $F$ , 30 units of  $C$ , and 20 units of  $T$ , the required units are 40, 30, and 20 respectively. Therefore, to produce 1 unit of each, we divide by 250:

$$\text{For } F: \frac{40}{250}, \text{ For } C: \frac{30}{250}, \text{ For } T: \frac{20}{250}$$

Similarly, for 220 units of  $C$ , the required units are 50, 40, and 30 respectively. To produce 1 unit of  $C$ , we divide by 220:

$$\text{For } F: \frac{50}{220}, \text{ For } C: \frac{40}{220}, \text{ For } T: \frac{30}{220}$$

And for 200 units of  $T$ , the required units are 60, 40, and 30 respectively. To produce 1 unit of  $T$ , we divide by 200:

$$\text{For } F: \frac{60}{200}, \text{ For } C: \frac{40}{200}, \text{ For } T: \frac{30}{200}$$

These fractions represent the units of  $F$ ,  $C$ , and  $T$  required to produce 1 unit of each.

We obtain the following matrix:

$$A = \begin{bmatrix} \frac{40}{250} & \frac{50}{220} & \frac{60}{200} \\ \frac{30}{250} & \frac{40}{220} & \frac{40}{200} \\ \frac{20}{250} & \frac{30}{220} & \frac{30}{200} \end{bmatrix} = \begin{bmatrix} .1600 & .2273 & .3000 \\ .1200 & .1818 & .2000 \\ .0800 & .1364 & .1500 \end{bmatrix}$$

Clearly  $AX + D = X$

$$\begin{bmatrix} .1600 & .2273 & .3000 \\ .1200 & .1818 & .2000 \\ .0800 & .1364 & .1500 \end{bmatrix} \begin{bmatrix} 250 \\ 220 \\ 200 \end{bmatrix} + \begin{bmatrix} 100 \\ 110 \\ 120 \end{bmatrix} = \begin{bmatrix} 250 \\ 220 \\ 200 \end{bmatrix}$$

**Summary 2.6.1: Leontief's Models****Leontief's Closed Model**

1. All consumption is within the industries. There is no external demand.
2. Input equals output.
3.  $X = AX$  or  $(I - A)X = 0$

**Leontief's Open Model**

1. In addition to internal consumption, there is an outside demand by the consumer.
2. Input equals output.
3.  $X = AX + D$  or  $X = (I - A)^{-1}D$



# Chapter 3

## Linear Programming with Geometry

In this chapter, you will learn to:

1. Solve linear programming problems that maximize the objective function.
2. Solve linear programming problems that minimize the objective function.

### 3.1 Maximization Applications

In this section, you will learn to:

1. Recognize the typical form of a linear programming problem.
2. Formulate maximization linear programming problems.
3. Graph feasibility regions for maximization linear programming problems.
4. Determine optimal solutions for maximization linear programming problems.

Application problems in business, economics, and social and life sciences often ask us to make decisions on the basis of certain conditions. The con-

ditions or constraints often take the form of inequalities. In this section, we will begin to formulate, analyze, and solve such problems, at a simple level, to understand the many components of such a problem.

A typical linear programming problem consists of finding an extreme value of a linear function subject to certain constraints. We are either trying to maximize or minimize the value of this linear function, such as to maximize profit or revenue, or to minimize cost. That is why these linear programming problems are classified as maximization or minimization problems, or just optimization problems. The function we are trying to optimize is called an objective function, and the conditions that must be satisfied are called constraints.

A typical example is to maximize profit from producing several products, subject to limitations on materials or resources needed for producing these items; the problem requires us to determine the amount of each item produced. Another type of problem involves scheduling; we need to determine how much time to devote to each of several activities in order to maximize income from (or minimize cost of) these activities, subject to limitations on time and other resources available for each activity.

In this chapter, we will work with problems that involve only two variables, and therefore, can be solved by graphing.

In the next chapter, we'll learn an algorithm to find a solution numerically. That will provide us with a tool to solve problems with more than two variables. At that time, with a little more knowledge about linear programming, we'll also explore the many ways these techniques are used in business and wide variety of other fields.

We begin by solving a maximization problem.

**Example 3.1.1.** *Niki holds two part-time jobs, Job I and Job II. She never wants to work more than a total of 12 hours a week. She has determined that for every hour she works at Job I, she needs 2 hours of preparation time, and for every hour she works at Job II, she needs one hour of preparation time, and she cannot spend more than 16 hours for preparation. If Niki makes \$40 an hour at Job I, and \$30 an hour at Job II, how many hours should she work per week at each job to maximize her income?*

**Solution 3.1.1.** *We start by choosing our variables. Let  $x$  be the number of*

hours per week Niki will work at Job I, and  $y$  the number of hours per week she will work at Job II.

Now we write the objective function. Since Niki gets paid \$40 an hour at Job I, and \$30 an hour at Job II, her total income  $I$  is given by the following equation.

$$I = 40x + 30y$$

Our next task is to find the constraints. The constraints based on the problem description are:

$$x + y \leq 12$$

$$2x + y \leq 16$$

$$x \geq 0, \quad y \geq 0$$

We have formulated the problem as follows: Maximize

$$I = 40x + 30y$$

Subject to:

$$x + y \leq 12$$

$$2x + y \leq 16$$

$$x \geq 0; \quad y \geq 0$$

To solve the problem, we graph the constraints and shade the region that satisfies all the inequality constraints. We graph the lines by plotting the  $x$ -intercept and  $y$ -intercept and use a test point to determine which portion of the plane to shade.

In this example, after graphing the lines representing the constraints and using the origin  $(0,0)$  as a test point, we find that the feasible region is the area below and to the left of both constraint lines, above the  $x$ -axis, and to the right of the  $y$ -axis.

The shaded region where all conditions are satisfied is called the feasibility region or the feasibility polygon. The Fundamental Theorem of Linear Programming states that the maximum (or minimum) value of the objective function always takes place at the vertices of the feasibility region. Therefore, we will identify all the vertices (corner points) of the feasibility region. We

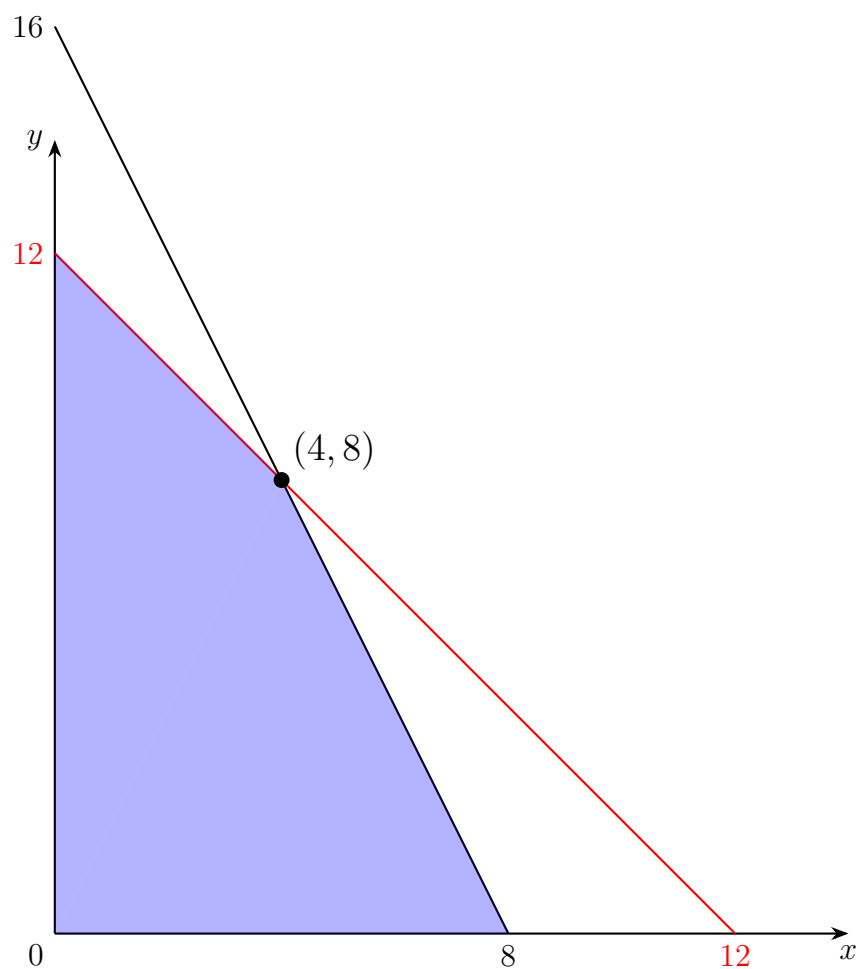


Figure 3.1: The red line is  $x + y = 12$ , the black line is  $2x + y = 16$ , and the blue region is the feasible region.

call these points critical points. They are listed as  $(0, 0)$ ,  $(0, 12)$ ,  $(4, 8)$ , and  $(8, 0)$ .

To maximize Niki's income, we will substitute these points in the objective function to see which point gives us the highest income per week. We list the results below:

Critical Points	Income
(0, 0)	$40(0) + 30(0) = \$0$
(0, 12)	$40(0) + 30(12) = \$360$
(4, 8)	$40(4) + 30(8) = \$400$
(8, 0)	$40(8) + 30(0) = \$320$

Clearly, the point (4, 8) gives the most profit: \$400. Therefore, we conclude that Niki should work 4 hours at Job I and 8 hours at Job II.

**Example 3.1.2.** *A factory manufactures two types of gadgets, regular and premium. Each gadget requires the use of two operations, assembly and finishing, and there are at most 12 hours available for each operation. A regular gadget requires 1 hour of assembly and 2 hours of finishing, while a premium gadget needs 2 hours of assembly and 1 hour of finishing. Due to other restrictions, the company can make at most 7 gadgets a day. If a profit of \$20 is realized for each regular gadget and \$30 for a premium gadget, how many of each should be manufactured to maximize profit?*

**Solution 3.1.2.** *We choose our variables. Let  $x$  be the number of regular gadgets manufactured each day, and  $y$  be the number of premium gadgets manufactured each day.*

*The objective function is*

$$P = 20x + 30y$$

*We now write the constraints. The company can make at most 7 gadgets a day, giving us:*

$$x + y \leq 7$$

*The regular gadget requires one hour of assembly and the premium gadget two hours, with at most 12 hours available for assembly:*

$$x + 2y \leq 12$$

*Similarly, for finishing, we have:*

$$2x + y \leq 12$$

*The non-negativity constraints are:*

$$x \geq 0, \quad y \geq 0$$

We formulate the problem as follows: Maximize  $P = 20x + 30y$  Subject to:

$$x + y \leq 7$$

$$x + 2y \leq 12$$

$$2x + y \leq 12$$

$$x \geq 0; \quad y \geq 0$$

We next graph the constraints and feasibility region.

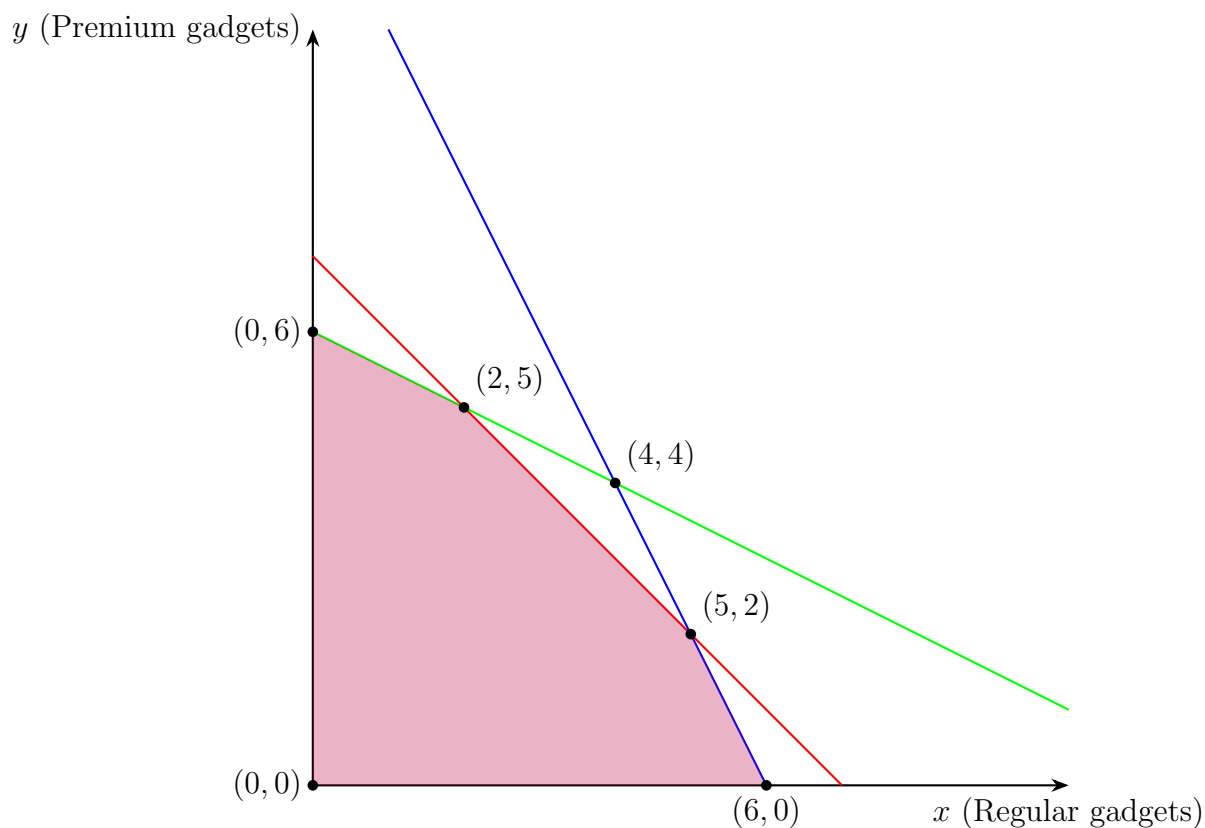


Figure 3.2: Feasibility region for the gadget factory optimization problem

Again, we have shaded the feasibility region, where all constraints are satisfied. Since the extreme value of the objective function always takes place at the vertices of the feasibility region, we identify all the critical points. They

are listed as  $(0, 0)$ ,  $(0, 6)$ ,  $(2, 5)$ ,  $(5, 2)$ , and  $(6, 0)$ . Notice,  $(4, 4)$  is **not** a critical point because it is not on the edge of the critical region. To maximize profit, we will substitute these points in the objective function to see which point gives us the maximum profit each day. The results are listed below:

Critical Point	Income
$(0, 0)$	$20(0) + 30(0) = \$0$
$(0, 6)$	$20(0) + 30(6) = \$180$
$(2, 5)$	$20(2) + 30(5) = \$190$
$(5, 2)$	$20(5) + 30(2) = \$160$
$(6, 0)$	$20(6) + 30(0) = \$120$

The point  $(2, 5)$  gives the most profit, and that profit is \$190. Therefore, we conclude that we should manufacture 2 regular gadgets and 5 premium gadgets daily to obtain the maximum profit of \$190.

So far, we have focused on "standard maximization problems" in which:

1. The objective function is to be maximized.
2. All constraints are of the form  $ax + by \leq c$ .
3. All variables are constrained to be non-negative ( $x \geq 0$ ,  $y \geq 0$ ).

We will next consider an example where that is not the case. Our next problem is said to have "mixed constraints" since some of the inequality constraints are of the form  $ax + by \leq c$  and some are of the form  $ax + by \geq c$ . The non-negativity constraints are still an important requirement in any linear program.

**Example 3.1.3.** *Solve the following maximization problem graphically.*

$$\begin{aligned}
 &\text{Maximize } P = 10x + 15y \\
 &\text{Subject to: } x + y \geq 1 \\
 &\quad \quad \quad x + 2y \leq 6 \\
 &\quad \quad \quad 2x + y \leq 6 \\
 &\quad \quad \quad x \geq 0; \quad y \geq 0
 \end{aligned}$$

**Solution 3.1.3.** *The graph is shown below.*

*The five critical points are listed in the figure above. The reader should observe that the first constraint  $x + y \geq 1$  requires that the feasibility region*

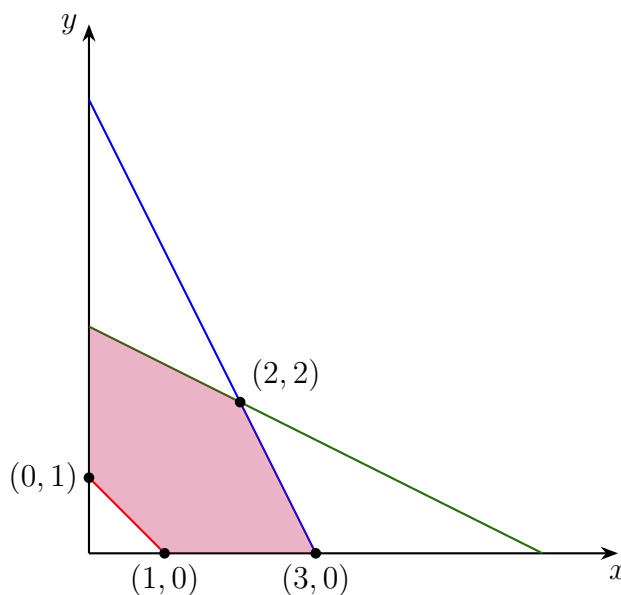


Figure 3.3: The red line is  $x + y = 1$ , the green line is  $x + 2y = 6$  and the blue line is  $2x + y = 6$ .

*must be bounded below by the line  $x + y = 1$ ; the test point  $(0,0)$  does not satisfy  $x + y \geq 1$ , so we shade the region on the opposite side of the line from the test point  $(0,0)$ .*

<i>Critical Point</i>	<i>Income</i>
$(1, 0)$	$10(1) + 15(0) = \$10$
$(3, 0)$	$10(3) + 15(0) = \$30$
$(2, 2)$	$10(2) + 15(2) = \$50$
$(0, 3)$	$10(0) + 15(3) = \$45$
$(0, 1)$	$10(0) + 15(1) = \$15$

*Clearly, the point  $(2, 2)$  maximizes the objective function to a maximum value of 50. It is important to observe that if the point  $(0,0)$  lies on the line for a constraint, then  $(0,0)$  could not be used as a test point. We would need to select any other point that does not lie on the line to use as a test point in that situation.*

Finally, we address an important question: Is it possible to determine the point that gives the maximum value without calculating the value at each



critical point?

The answer is yes.

For example 3.1.2, we substituted the points  $(0, 0)$ ,  $(0, 6)$ ,  $(2, 5)$ ,  $(5, 2)$ , and  $(6, 0)$  in the objective function  $P = 20x + 30y$ , and we got the values \$0, \$180, \$190, \$160, \$120, respectively. Sometimes that is not the most efficient way of finding the optimum solution. Instead, we could find the optimal value by also graphing the objective function.

To determine the largest  $P$ , we graph  $P = 20x + 30y$  for any value  $P$  of our choice. Let us say, we choose  $P = 60$ . We graph  $20x + 30y = 60$ .

Now we move the line parallel to itself, that is, keeping the same slope at all times. Since we are moving the line parallel to itself, the slope is kept the same, and the only thing that is changing is the  $P$ . As we move away from the origin, the value of  $P$  increases. The largest possible value of  $P$  is realized when the line touches the last corner point of the feasibility region.

The figure below shows the movements of the line, and the optimum solution is achieved at the point  $(2, 5)$ . In maximization problems, as the line is being moved away from the origin, this optimum point is the farthest critical point.

#### Summary 3.1.1: Maximization Linear Programming Problems

1. Write the objective function.
2. Write the constraints.
  - (a) For the standard maximization linear programming problems, constraints are of the form:  $ax + by \leq c$ .
  - (b) Since the variables are non-negative, we include the constraints:  $x \geq 0, y \geq 0$ .
3. Graph the constraints.
4. Shade the feasibility region.
5. Find the corner points.
6. Determine the corner point that gives the maximum value.
  - (a) This is done by finding the value of the objective function at each corner point.
  - (b) This can also be done by moving the line associated with the objective function.

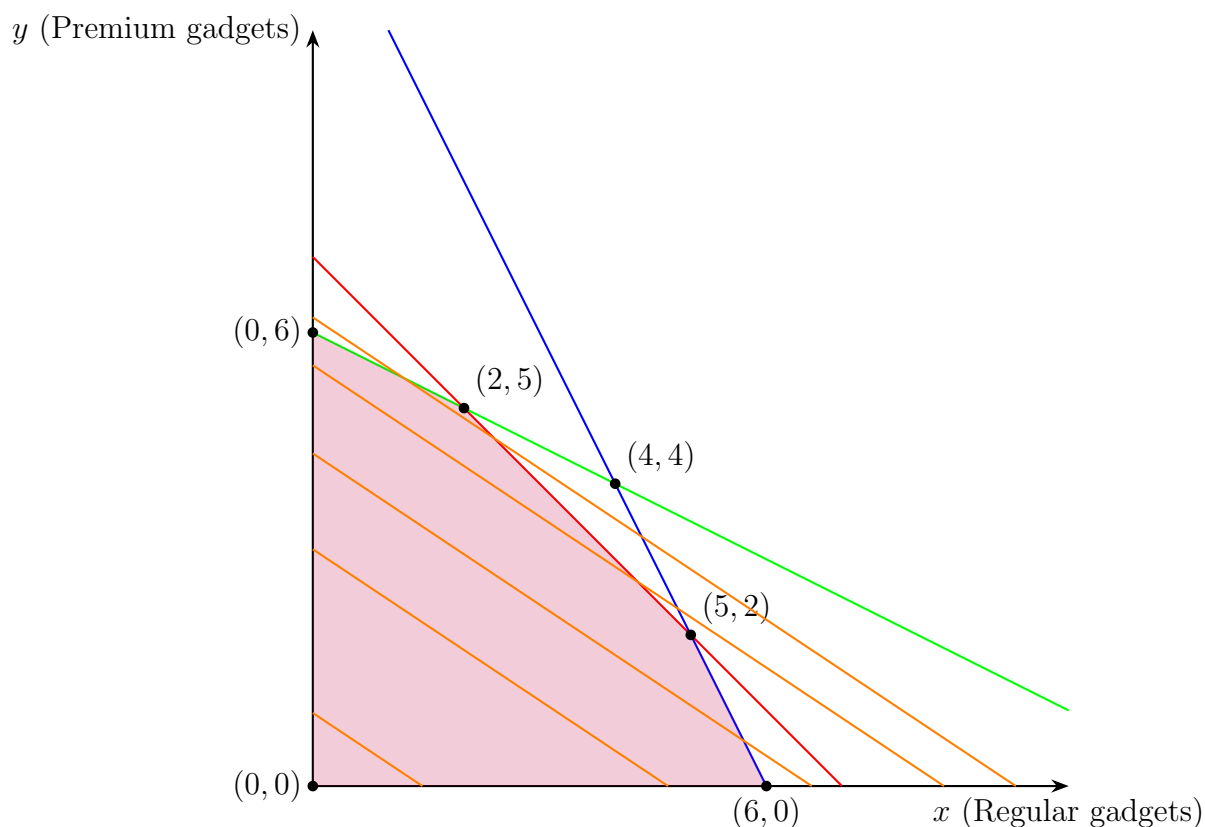


Figure 3.4: Feasibility region for the gadget factory optimization problem with profit lines.

## 3.2 Minimization Applications

In this section, you will learn to:

1. Formulate minimization linear programming problems.
2. Graph feasibility regions for minimization linear programming problems.
3. Determine optimal solutions for minimization linear programming problems.

Minimization linear programming problems are solved in much the same way as the maximization problems.

For the standard minimization linear program, the constraints are of the form  $ax + by \geq c$ , as opposed to the form  $ax + by \leq c$  for the standard maximization problem. As a result, the feasible solution extends indefinitely to the upper right of the first quadrant, and is unbounded. But that is not a concern, since in order to minimize the objective function, the line associated with the objective function is moved towards the origin, and the critical point that minimizes the function is closest to the origin.

However, one should be aware that in the case of an unbounded feasibility region, the possibility of no optimal solution exists.

**Example 3.2.1.** *At a university, Professor Symons wishes to employ two people, John and Mary, to grade papers for his classes. John is a graduate student and can grade 20 papers per hour; John earns \$15 per hour for grading papers. Mary is a post-doctoral associate and can grade 30 papers per hour; Mary earns \$25 per hour for grading papers. Each must be employed at least one hour a week to justify their employment. If Professor Symons has at least 110 papers to be graded each week, how many hours per week should he employ each person to minimize the cost?*

**Solution 3.2.1.** *We choose the variables as follows: Let  $x$  be the number of hours per week John is employed, and  $y$  be the number of hours per week Mary is employed.*

*The objective function is*

$$C = 15x + 25y$$

*The constraints are that each must work at least one hour each week:*

$$x \geq 1$$

$$y \geq 1$$

*John can grade 20 papers per hour and Mary 30 papers per hour, with at least 110 papers to be graded per week:*

$$20x + 30y \geq 110$$

*Additionally,  $x$  and  $y$  are non-negative:*

$$x \geq 0$$

$$y \geq 0$$

The problem is thus formulated as: Minimize  $C = 15x + 25y$  Subject to:

$$x \geq 1$$

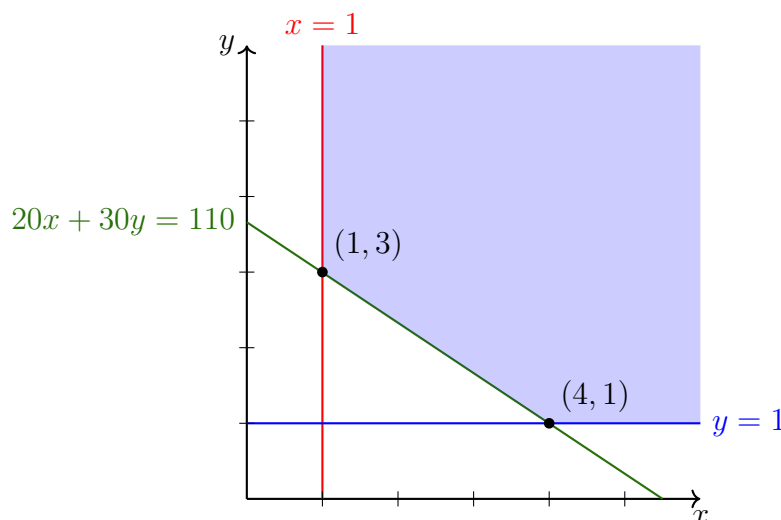
$$y \geq 1$$

$$20x + 30y \geq 110$$

$$x \geq 0$$

$$y \geq 0$$

To solve the problem, we graph the constraints as follows:



Again, we have shaded the feasibility region, where all constraints are satisfied. If we used test point  $(0,0)$  that does not lie on any of the constraints, we observe that  $(0,0)$  does not satisfy any of the constraints  $x \geq 1$ ,  $y \geq 1$ , and  $20x + 30y \geq 110$ . Thus, all the shading for the feasibility region lies on the opposite side of the constraint lines from the point  $(0,0)$ .

Alternatively, we could use test point  $(4,6)$ , which also does not lie on any of the constraint lines. We'd find that  $(4,6)$  does satisfy all of the inequality constraints. Consequently, all the shading for the feasibility region lies on the same side of the constraint lines as the point  $(4,6)$ .

Since the extreme value of the objective function always takes place at the vertices of the feasibility region, we identify the two critical points,  $(1, 3)$  and  $(4, 1)$ . To minimize cost, we will substitute these points in the objective function to see which point gives us the minimum cost each week. The results are listed below:

<i>Critical points</i>	<i>Income</i>
$(1, 3)$	$15(1) + 25(3) = \$90$
$(4, 1)$	$15(4) + 25(1) = \$85$

The point  $(4, 1)$  gives the least cost, and that cost is \$85. Therefore, we conclude that in order to minimize grading costs, Professor Symons should employ John for 4 hours a week and Mary for 1 hour a week at a cost of \$85 per week.

**Example 3.2.2.** Professor Hamer is on a low cholesterol diet. During lunch at the college cafeteria, he always chooses between two meals, Pasta or Tofu. The table below lists the amount of protein, carbohydrates, and vitamins each meal provides along with the amount of cholesterol he is trying to minimize. Mr. Hamer needs at least 200 grams of protein, 960 grams of carbohydrates, and 40 grams of vitamins for lunch each month. Over this time period, how many days should he have the Pasta meal, and how many days the Tofu meal so that he gets the adequate amount of protein, carbohydrates, and vitamins and at the same time minimizes his cholesterol intake?

	<b><i>Pasta</i></b>	<b><i>Tofu</i></b>
<i>Protein (g)</i>	8	16
<i>Carbohydrates (g)</i>	60	40
<i>Vitamin C (g)</i>	2	2
<i>Cholesterol (mg)</i>	60	50

**Solution 3.2.2.** We choose the variables as follows: Let  $x$  be the number of days Mr. Hamer eats Pasta, and  $y$  the number of days he eats Tofu.

The objective function for minimizing cholesterol intake is

$$C = 60x + 50y$$

*The constraints for protein, carbohydrates, and vitamins are as follows:*

$$\begin{aligned}8x + 16y &\geq 200 \\60x + 40y &\geq 960 \\2x + 2y &\geq 40\end{aligned}$$

*Additionally,  $x$  and  $y$  are non-negative:*

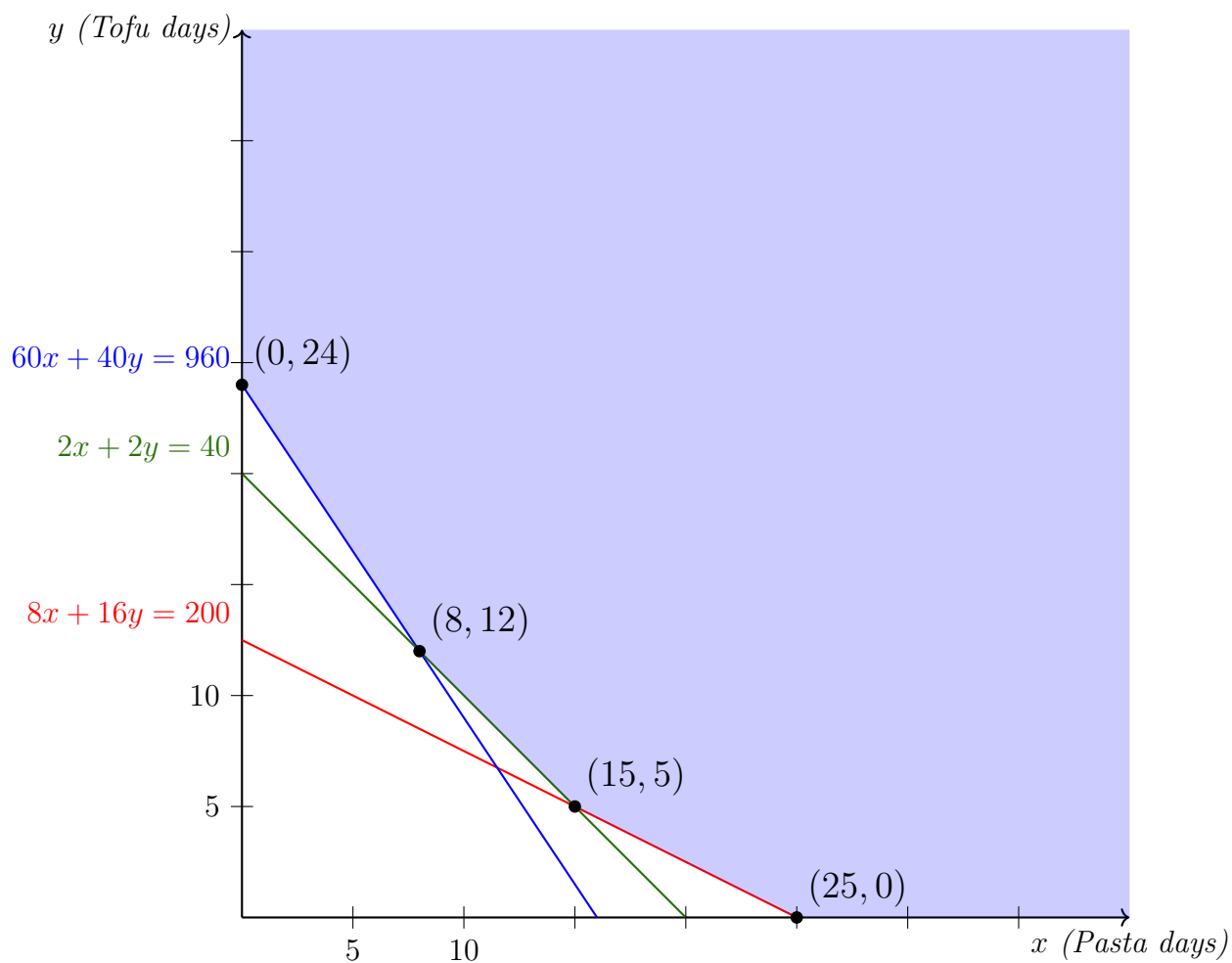
$$x \geq 0$$

$$y \geq 0$$

*We summarize the problem as: Minimize  $C = 60x + 50y$  Subject to:*

$$\begin{aligned}8x + 16y &\geq 200 \\60x + 40y &\geq 960 \\2x + 2y &\geq 40 \\x &\geq 0 \\y &\geq 0\end{aligned}$$

*To solve the problem, we graph the constraints and shade the feasibility region.*



We have shaded the unbounded feasibility region, where all constraints are satisfied. To minimize the objective function, we find the vertices of the feasibility region. These vertices are  $(0, 24)$ ,  $(8, 12)$ ,  $(15, 5)$ , and  $(25, 0)$ . To minimize cholesterol, we will substitute these points in the objective function to see which point gives us the smallest value. The results are listed below:

Critical points	Cholesterol
$(0, 24)$	$60(0) + 50(24) = 1200 \text{ mg}$
$(8, 12)$	$60(8) + 50(12) = 1080 \text{ mg}$
$(15, 5)$	$60(15) + 50(5) = 1150 \text{ mg}$
$(25, 0)$	$60(25) + 50(0) = 1500 \text{ mg}$

*The point  $(8, 12)$  gives the least cholesterol, which is 1080 mg. This states that for every 20 meals, Professor Hamer should eat Pasta for 8 days and Tofu for 12 days.*

We must be aware that in some cases, a linear program may not have an optimal solution.

- A linear program can fail to have an optimal solution if there is no feasibility region. If the inequality constraints are not compatible, there may not be a region in the graph that satisfies all the constraints. If the linear program does not have a feasible solution satisfying all constraints, then it cannot have an optimal solution.
- A linear program can fail to have an optimal solution if the feasibility region is unbounded. The two minimization linear programs we examined had unbounded feasibility regions. The feasibility region was bounded by constraints on some sides but was not entirely enclosed by the constraints. Both of the minimization problems had optimal solutions. However, if we were to consider a maximization problem with a similar unbounded feasibility region, the linear program would have no optimal solution. No matter what values of  $x$  and  $y$  were selected, we could always find other values of  $x$  and  $y$  that would produce a higher value for the objective function. In other words, if the value of the objective function can be increased without bound in a linear program with an unbounded feasible region, there is no optimal maximum solution.

Although the method of solving minimization problems is similar to that of maximization problems, we still feel that we should summarize the steps involved.



**Summary 3.2.1: Minimization Linear Programming Problems**

1. Write the objective function.
2. Write the constraints.
  - (a) For standard minimization linear programming problems, constraints are of the form:  $ax + by \geq c$ .
  - (b) Since the variables are non-negative, include the constraints:  $x \geq 0$ ;  $y \geq 0$ .
3. Graph the constraints.
4. Shade the feasibility region.
5. Find the corner points.
6. Determine the corner point that gives the minimum value.
  - (a) This can be done by finding the value of the objective function at each corner point.
  - (b) This can also be done by moving the line associated with the objective function.
  - (c) There is the possibility that the problem has no solution.



## Chapter 4

# Linear Programming, Simplex Method

In this chapter, you will learn to:

1. Investigate real world applications of linear programming and related methods.
2. Solve linear programming maximization problems using the simplex method.
3. Solve linear programming minimumization problems using the simplex method.

### 4.1 Linear Programming Applications in Business, Finance, Medicine, and Social Science

In this section, you will learn about:

1. real world applications of linear programming and related methods

The linear programs we solved in chapter 3 contain only two variables,  $x$  and  $y$ , so that we could solve them graphically. In practice, linear programs can contain thousands of variables and constraints.

Later in this chapter we'll learn to solve linear programs with more than two variables using the simplex algorithm, which is a numerical solution method that uses matrices and row operations. However, in order to make the problems practical for learning purposes, our problems will still have only several variables.

Now that we understand the main concepts behind linear programming, we can also consider how linear programming is currently used in large scale real-world applications.

Linear programming is used in business and industry in production planning, transportation and routing, and various types of scheduling. Airlines use linear programs to schedule their flights, taking into account both scheduling aircraft and scheduling staff. Delivery services use linear programs to schedule and route shipments to minimize shipment time or minimize cost. Retailers use linear programs to determine how to order products from manufacturers and organize deliveries with their stores. Manufacturing companies use linear programming to plan and schedule production. Financial institutions use linear programming to determine the mix of financial products they offer, or to schedule payments transferring funds between institutions. Health care institutions use linear programming to ensure the proper supplies are available when needed. And as we'll see below, linear programming has also been used to organize and coordinate life saving health care procedures.

In some of the applications, the techniques used are related to linear programming but are more sophisticated than the methods we study in this class. One such technique is called integer programming. In these situations, answers must be integers to make sense, and can not be fractions. Problems where solutions must be integers are more difficult to solve than the linear programs we've worked with. In fact, many of our problems have been very carefully constructed for learning purposes so that the answers just happen to turn out to be integers, but in the real world unless we specify that as a restriction, there is no guarantee that a linear program will produce integer solutions. There are also related techniques that are called non-linear programs, where the functions defining the objective function and/or some or all of the constraints may be non-linear rather than straight lines.

Many large businesses that use linear programming and related methods have analysts on their staff who can perform the analyses needed, including linear programming and other mathematical techniques. Consulting firms

#### 4.1. LINEAR PROGRAMMING APPLICATIONS IN BUSINESS, FINANCE, MEDICINE, AND SOC

specializing in use of such techniques also aid businesses who need to apply these methods to their planning and scheduling processes.

When used in business, many different terms may be used to describe the use of techniques such as linear programming as part of mathematical business models. Optimization, operations research, business analytics, data science, industrial engineering and management science are among the terms used to describe mathematical modelling techniques that may include linear programming and related met

In the rest of this section we'll explore six real world applications, and investigate what they are trying to accomplish using optimization, as well as what their constraints might represent.

##### 4.1.1 Airline Scheduling

Airlines use techniques that include and are related to linear programming to schedule their aircraft to flights on various routes and to schedule crews to the flights. In addition, airlines also use linear programming to determine ticket pricing for various types of seats and levels of service or amenities, as well as the timing at which ticket prices change.

The process of scheduling aircraft and departure times on flight routes can be expressed as a model that minimizes cost, of which the largest component is generally fuel costs. Constraints involve considerations such as:

- Each aircraft needs to complete a daily or weekly tour to return back to its point of origin.
- Scheduling sufficient flights to meet demand on each route.
- Scheduling the right type and size of aircraft on each route to be appropriate for the route and for the demand for the number of passengers.
- Aircraft must be compatible with the airports it departs from and arrives at - not all airports can handle all types of planes.

A model to accomplish this could contain thousands of variables and constraints. Highly trained analysts determine ways to translate all the constraints into mathematical inequalities or equations to put into the model.

After aircraft are scheduled, crews need to be assigned to flights. Each flight

needs a pilot, a co-pilot, and flight attendants. Each crew member needs to complete a daily or weekly tour to return back to his or her home base. Additional constraints on flight crew assignments take into account factors such as:

- Pilot and co-pilot qualifications to fly the particular type of aircraft they are assigned to.
- Flight crew have restrictions on the maximum amount of flying time per day and the length of mandatory rest periods between flights or per day that must meet certain minimum rest time regulations.
- Numbers of crew members required for a particular type or size of aircraft.

When scheduling crews to flights, the objective function would seek to minimize total flight crew costs, determined by the number of people on the crew and pay rates of the crew members. However, the cost for any particular route might not end up being the lowest possible for that route, depending on tradeoffs to the total cost of shifting different crews to different routes.

An airline can also use linear programming to revise schedules on short notice on an emergency basis when there is a schedule disruption, such as due to weather. In this case, the considerations to be managed involve:

- Getting aircraft and crews back on schedule as quickly as possible.
- Moving aircraft from storm areas to areas with calm weather to keep the aircraft safe from damage and ready to come back into service as quickly and conveniently as possible.
- Ensuring crews are available to operate the aircraft and that crews continue to meet mandatory rest period requirements and regulations.

### **4.1.2 Kidney Donation Chain**

For patients who have kidney disease, a transplant of a healthy kidney from a living donor can often be a lifesaving procedure. Criteria for a kidney donation procedure include the availability of a donor who is healthy enough to donate a kidney, as well as a compatible match between the patient and donor for blood type and several other characteristics. Ideally, if a patient needs a kidney donation, a close relative may be a match and can be the

#### 4.1. LINEAR PROGRAMMING APPLICATIONS IN BUSINESS, FINANCE, MEDICINE, AND SOC

kidney donor. However, often there is not a relative who is a close enough match to be the donor. Considering donations from unrelated donors allows for a larger pool of potential donors. Kidney donations involving unrelated donors can sometimes be arranged through a chain of donations that pair patients with donors. For example, a kidney donation chain with three donors might operate as follows:

- Donor A donates a kidney to Patient B.
- Donor B, who is related to Patient B, donates a kidney to Patient C.
- Donor C, who is related to Patient C, donates a kidney to Patient A, who is related to Donor A.

Linear programming is one of several mathematical tools that have been used to help efficiently identify a kidney donation chain. In this type of model, patient/donor pairs are assigned compatibility scores based on characteristics of patients and potential donors. The objective is to maximize the total compatibility scores. Constraints ensure that donors and patients are paired only if compatibility scores are sufficiently high to indicate an acceptable match.

#### 4.1.3 Advertisements in Online Marketing

Did you ever make a purchase online and then notice that as you browse websites, search, or use social media, you now see more ads related the item you purchased? Marketing organizations use a variety of mathematical techniques, including linear programming, to determine individualized advertising placement purchases.

Instead of advertising randomly, online advertisers want to sell bundles of advertisements related to a particular product to batches of users who are more likely to purchase that product. Based on an individual's previous browsing and purchase selections, he or she is assigned a "propensity score" for making a purchase if shown an ad for a certain product. The company placing the ad generally does not know individual personal information based on the history of items viewed and purchased, but instead has aggregated information for groups of individuals based on what they view or purchase. However, the company may know more about an individual's history if he or she logged into a website making that information identifiable, within the

privacy provisions and terms of use of the site.

The company's goal is to buy ads to present to specified size batches of people who are browsing. The linear program would assign ads and batches of people to view the ads using an objective function that seeks to maximize advertising response modeled using the propensity scores. The constraints are to stay within the restrictions of the advertising budget.

#### **4.1.4 Loans**

A car manufacturer sells its cars through dealers. Dealers can offer loan financing to customers who need to take out loans to purchase a car. Here we will consider how car manufacturers can use linear programming to determine the specific characteristics of the loan they offer to a customer who purchases a car. In a future chapter, we will learn how to do the financial calculations related to loans.

A customer who applies for a car loan fills out an application. This provides the car dealer with information about that customer. In addition, the car dealer can access a credit bureau to obtain information about a customer's credit score.

Based on this information obtained about the customer, the car dealer offers a loan with certain characteristics, such as interest rate, loan amount, and length of loan repayment period.

Linear programming can be used as part of the process to determine the characteristics of the loan offer. The linear program seeks to maximize the profitability of its portfolio of loans. The constraints limit the risk that the customer will default and will not repay the loan. The constraints also seek to minimize the risk of losing the loan customer if the conditions of the loan are not favorable enough; otherwise, the customer may find another lender, such as a bank, which can offer a more favorable loan.

#### **4.1.5 Production Planning and Scheduling in Manufacturing**

Consider the example of a company that produces yogurt. There are different varieties of yogurt products in a variety of flavors. Yogurt products have a



#### 4.1. LINEAR PROGRAMMING APPLICATIONS IN BUSINESS, FINANCE, MEDICINE, AND SOC

short shelf life; it must be produced on a timely basis to meet demand, rather than drawing upon a stockpile of inventory as can be done with a product that is not perishable. Most ingredients in yogurt also have a short shelf life, so can not be ordered and stored for long periods of time before use; ingredients must be obtained in a timely manner to be available when needed but still be fresh. Linear programming can be used in both production planning and scheduling.

To start the process, sales forecasts are developed to determine demand to know how much of each type of product to make. There are often various manufacturing plants at which the products may be produced. The appropriate ingredients need to be at the production facility to produce the products assigned to that facility. Transportation costs must be considered, both for obtaining and delivering ingredients to the correct facilities, and for transport of finished product to the sellers. The linear program that monitors production planning and scheduling must be updated frequently - daily or even twice each day - to take into account variations from a master plan.

##### 4.1.6 Bike Share Programs

Over 600 cities worldwide have bikeshare programs. Although bikeshare programs have been around for a long time, they have proliferated in the past decade as technology has developed new methods for tracking the bicycles.

Bikeshare programs vary in the details of how they work, but most typically people pay a fee to join and then can borrow a bicycle from a bike share station and return the bike to the same or a different bike share station. Over time the bikes tend to migrate; there may be more people who want to pick up a bike at station A and return it at station B than there are people who want to do the opposite. In chapter 9, we'll investigate a technique that can be used to predict the distribution of bikes among the stations.

Once other methods are used to predict the actual and desired distributions of bikes among the stations, bikes may need to be transported between stations to even out the distribution. Bikeshare programs in large cities have used methods related to linear programming to help determine the best routes and methods for redistributing bicycles to the desired stations once the desired distributions have been determined. The optimization model would seek to minimize transport costs and/or time subject to constraints of having

sufficient bicycles at the various stations to meet demand.

## 4.2 Maximization by the Simplex Method

In this section, you will learn to:

1. Solve linear programming maximization problems using the Simplex Method by
  - (a) Identifying and set up a linear program in standard maximization form
  - (b) Converting inequality constraints to equations using slack variables
  - (c) Setting up the initial simplex tableau using the objective function and slack equations
  - (d) Finding the optimal simplex tableau by performing pivoting operations
  - (e) Identifying the optimal solution from the optimal simplex tableau

In the last chapter, we used the geometrical method to solve linear programming problems, but the geometrical approach will not work for problems that have more than two variables. In real-life situations, linear programming problems consist of literally thousands of variables and are solved by computers. We can solve these problems algebraically, but that will not be very efficient. Suppose we were given a problem with, say, 5 variables and 10 constraints. By choosing all combinations of five equations with five unknowns, we could find all the corner points, test them for feasibility, and come up with the solution, if it exists. But the trouble is that even for a problem with so few variables, we will get more than 250 corner points, and testing each point will be very tedious. So we need a method that has a systematic algorithm and can be programmed for a computer. The method has to be efficient enough so we wouldn't have to evaluate the objective function at each corner point. We have just such a method, and it is called the simplex method.

The simplex method was developed during the Second World War by Dr. George Dantzig. His linear programming models helped the Allied forces

with transportation and scheduling problems. In 1979, a Soviet scientist named Leonid Khachian developed a method called the ellipsoid algorithm, which was supposed to be revolutionary, but as it turned out, it is not any better than the simplex method. In 1984, Narendra Karmarkar, a research scientist at AT&T Bell Laboratories developed Karmarkar's algorithm, which has been proven to be four times faster than the simplex method for certain problems. But the simplex method still works the best for most problems.

The simplex method uses an approach that is very efficient. It does not compute the value of the objective function at every point; instead, it begins with a corner point of the feasibility region where all the main variables are zero and then systematically moves from corner point to corner point while improving the value of the objective function at each stage. The process continues until the optimal solution is found.

To learn the simplex method, we try a rather unconventional approach. We first list the algorithm, and then work a problem. We justify the reasoning behind each step during the process. A thorough justification is beyond the scope of this course.

We start out with an example we solved in the last chapter by the graphical method. This will provide us with some insight into the simplex method and at the same time give us the chance to compare a few of the feasible solutions we obtained previously by the graphical method.

But first, we list the algorithm for the simplex method.

**Summary 4.2.1: The Simplex Method**

Here are the steps to solve a linear programming problem using the Simplex Method:

1. **Set up the problem.**
  - Write the objective function and the inequality constraints.
2. **Convert the inequalities into equations.**
  - Add one slack variable for each inequality.
3. **Construct the initial simplex tableau.**
  - Write the objective function as the bottom row.
4. **Identify the pivot column.**
  - The most negative entry in the bottom row identifies the pivot column.
5. **Calculate the quotients.**
  - Divide the far-right column by the identified pivot column to find quotients.
  - The smallest positive quotient identifies a row, and its corresponding element is the pivot element.
6. **Perform pivoting.**
  - Make all other entries in the pivot column zero by using the Gauss-Jordan method.
7. **Repeat if necessary.**
  - If there are still negative entries in the bottom row, go back to step 4.
8. **Read off your answers.**
  - Get the variables using the columns with 1 and 0s. All other variables are zero.
  - The maximum value you are looking for appears in the bottom right-hand corner.

Now, we use the simplex method to solve example 4.2.1 solved geometrically in example 3.1.1.

**Example 4.2.1.** *Niki holds two part-time jobs, Job I and Job II. She never wants to work more than a total of 12 hours a week. For every hour she works at Job I, she needs 2 hours of preparation time, and for every hour at Job II, she needs one hour of preparation time. She cannot spend more than 16 hours on preparation. If Niki makes \$40 an hour at Job I and \$30 an hour at Job II, how many hours should she work at each job to maximize*

her income?

**Solution 4.2.1.**

1. **Set up the problem.** Write the objective function and the constraints. Since the simplex method is used for problems that consist of many variables, it is not practical to use the variables  $x, y, z$  etc. We use symbols  $x_1, x_2, x_3$ , and so on.

Let  $x_1$  = The number of hours per week Niki will work at Job I.  
and  $x_2$  = The number of hours per week Niki will work at Job II.

It is customary to choose the variable that is to be maximized as  $Z$ . The problem is formulated the same way as we did in the last chapter.

$$\text{Maximize } Z = 40x_1 + 30x_2$$

Subject to:

$$\begin{aligned} x_1 + x_2 &\leq 12 \\ 2x_1 + x_2 &\leq 16 \\ x_1, x_2 &\geq 0 \end{aligned}$$

2. **Convert the inequalities into equations.** This is done by adding one slack variable for each inequality.

For example, to convert the inequality  $x_1 + x_2 \leq 12$  into an equation, we add a non-negative variable  $y_1$ , and we get

$$x_1 + x_2 + y_1 = 12$$

Here the variable  $y_1$  picks up the slack, and it represents the amount by which  $x_1 + x_2$  falls short of 12. In this problem, if Niki works fewer than 12 hours, say 10, then  $y_1$  is 2. Later when we read off the final solution from the simplex table, the values of the slack variables will identify the unused amounts.

We rewrite the objective function  $Z = 40x_1 + 30x_2$  as  $-40x_1 - 30x_2 + Z = 0$ .

After adding the slack variables, our problem reads

Objective function:

$$-40x_1 - 30x_2 + Z = 0$$

Subject to constraints:

$$x_1 + x_2 + y_1 = 12$$

$$2x_1 + x_2 + y_2 = 16$$

$$x_1, x_2 \geq 0$$

3. **Construct the initial simplex tableau.** Each inequality constraint appears in its own row. (The non-negativity constraints do not appear as rows in the simplex tableau.) Write the objective function as the bottom row.

Now that the inequalities are converted into equations, we can represent the problem into an augmented matrix called the initial simplex tableau as follows.

$x_1$	$x_2$	$y_1$	$y_2$	$Z$	$C$
1	1	1	0	0	12
2	1	0	1	0	16
-40	-30	0	0	1	0

Here the vertical line separates the left hand side of the equations from the right side. The horizontal line separates the constraints from the objective function. The right side of the equation is represented by the column  $C$ .

The reader may observe that the last four columns of this matrix look like the final matrix for the solution of a system of equations. If we arbitrarily choose  $x_1 = 0$  and  $x_2 = 0$ , we get

$y_1$	$y_2$	$Z$	$C$
1	0	0	12
0	1	0	16
0	0	1	0

which reads  $y_1 = 12$ ,  $y_2 = 16$ ,  $Z = 0$ .

The solution obtained by arbitrarily assigning values to some variables and then solving for the remaining variables is called the basic solution associated with the tableau. So the above solution is the basic solution associated with the initial simplex tableau. We can label the basic solution variable in the right of the last column as shown in the table below.

$x_1$	$x_2$	$y_1$	$y_2$	$Z$	$C$	
1	1	1	0	0	12	$y_1$
2	1	0	1	0	16	$y_2$
-40	-30	0	0	1	0	$Z$

4. **The most negative entry in the bottom row identifies the pivot column.** The most negative entry in the bottom row is -40; therefore, the column 1 is identified.

$x_1$	$x_2$	$y_1$	$y_2$	$Z$	$C$
1	1	1	0	0	12
2	1	0	1	0	16
-40	-30	0	0	1	0

**Why do we choose the most negative entry in the bottom row?**

The most negative entry in the bottom row represents the largest coefficient in the objective function; the coefficient whose entry will increase the value of the objective function the quickest.

The simplex method begins at a corner point where all the main variables, the variables that have symbols such as  $x_1$ ,  $x_2$ ,  $x_3$ , etc., are zero. It then moves from a corner point to the adjacent corner point always increasing the value of the objective function. In the case of the objective function  $Z = 40x_1 + 30x_2$ , it will make more sense to increase the value of  $x_1$  rather than  $x_2$ . The variable  $x_1$  represents the number of hours per week Niki works at Job I. Since Job I pays \$40 per hour as opposed to Job II which pays only \$30, the variable  $x_1$  will increase the objective function by \$40 for a unit of increase in the variable  $x_1$ .

5. **Calculate the quotients. The smallest quotient identifies a row.** The element in the intersection of the column identified in step 4 (marked with  $\uparrow$ ) and the row identified in this step is identified as the pivot element.

Following the algorithm, in order to calculate the quotient, we divide the entries in the far right column by the entries in column 1, excluding the entry in the bottom row.

$x_1$	$x_2$	$y_1$	$y_2$	$Z$	$C$	
1	1	1	0	0	12	$12/1 = 12$
2	1	0	1	0	16	$\leftarrow 16/2 = 8$
-40	-30	0	0	1	0	
$\uparrow$						

The smallest of the two quotients, 12 and 8, is 8. Therefore row 2 is identified. The intersection of column 1 and row 2 is the entry 2, which has been highlighted. This is our pivot element.

**Why do we find quotients, and why does the smallest quotient identify a row?**

When we choose the most negative entry in the bottom row, we are trying to increase the value of the objective function by bringing in the variable  $x_1$ . But we cannot choose any value for  $x_1$ . For instance, letting  $x_1 = 100$  is not possible because Niki never wants to work more than 12 hours at both jobs combined:  $x_1 + x_2 \leq 12$ . Therefore, the maximum she can work is 12 hours for  $x_1$ , meaning the preparation time for Job I is two times the time spent on the job. Since she never wants to spend more than 16 hours for preparation, the maximum time she can work is  $\frac{16}{2} = 8$  hours. Using the pivot element guarantees that we do not violate the constraints.

**Why do we identify the pivot element?**

The simplex method begins at a corner point where all the main variables, the variables that have symbols such as  $x_1, x_2, x_3$ , etc., are zero. It then moves from a corner point to the adjacent corner point always improving the value of the objective function. The value of the objective function is improved by changing the number of units of the variables.



We may add the number of units of one variable, while throwing away the units of another. Pivoting allows us to do just that.

The variable whose units are being added is called the entering variable, and the variable whose units are being replaced is called the departing variable. The entering variable in the above table is  $x_1$ , and it was identified by the most negative entry in the bottom row. The departing variable  $y_2$  was identified by the lowest of all quotients.

6. **Perform pivoting to make all other entries in this column zero.** In Chapter 2, we used pivoting to obtain the row echelon form of an augmented matrix. Pivoting is a process of obtaining a 1 in the location of the pivot element (marked by a box below), and then making all other entries zeros in that column. We've highlighted the pivot row to make it easier to track. So now our job is to make our pivot element a 1 by dividing the entire second row by 2. The result follows.

$x_1$	$x_2$	$y_1$	$y_2$	$Z$	$C$
1	1	1	0	0	12
<span style="border: 1px solid black;">2</span>	1	0	1	0	16
-40	-30	0	0	1	0

To obtain a zero in the entry first above the pivot element, we multiply the second row by  $-1$  and add it to row 1. We get

$x_1$	$x_2$	$y_1$	$y_2$	$Z$	$C$
0	$1/2$	1	$-1/2$	0	4
<span style="border: 1px solid black;">1</span>	$1/2$	0	$1/2$	0	8
-40	-30	0	0	1	0

To obtain a zero in the element below the pivot, we multiply the second row by 40 and add it to the last row.

$x_1$	$x_2$	$y_1$	$y_2$	$Z$	$C$
0	$1/2$	1	$-1/2$	0	4
<span style="border: 1px solid black;">1</span>	$1/2$	0	$1/2$	0	8
0	-10	0	20	1	320

We now determine the basic solution associated with this tableau. By arbitrarily choosing  $x_2 = 0$  and  $y_2 = 0$ , we obtain  $x_1 = 8$ ,  $y_1 = 4$ , and  $Z = 320$ . If we write the augmented matrix, whose left side is a matrix with columns that have one 1 and all other entries zeros, we get the following matrix stating the same thing.

$$\left[ \begin{array}{ccc|c} x_1 & y_1 & Z & C \\ 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & 320 \end{array} \right]$$

We can restate the solution associated with this matrix as  $x_1 = 8$ ,  $x_2 = 0$ ,  $y_1 = 4$ ,  $y_2 = 0$ , and  $z = 320$ . At this stage, it reads that if Niki works 8 hours at Job I and no hours at Job II, her profit  $z$  will be \$320. Recall from Example 1 in Section 3.1 that  $(8, 0)$  was one of our corner points. Here  $y_1 = 4$  and  $y_2 = 0$  mean that she will be left with 4 hours of working time and no preparation time.

7. **When there are no more negative entries in the bottom row, we are finished; otherwise, we start again from step 4.** Since there is still a negative entry,  $-10$ , in the bottom row, we need to begin, again, from step 4. This time we will not repeat the details of every step; instead, we will identify the column and row that give us the pivot element, and highlight the pivot element. The result is as follows.

$$\begin{array}{ccccc|c} x_1 & x_2 & y_1 & y_2 & Z & C \\ 0 & 1/2 & 1 & -1/2 & 0 & 4 \leftarrow 4/(1/2) = 8 \\ 1 & 1/2 & 0 & 1/2 & 0 & 8 \quad 8/(1/2) = 16 \\ \hline 0 & -10 & 0 & 20 & 1 & 320 \\ & \uparrow & & & & \end{array}$$

We make the pivot element 1 by multiplying row 1 by 2, and we get

$$\begin{array}{ccccc|c} x_1 & x_2 & y_1 & y_2 & Z & C \\ 0 & \boxed{1} & 2 & -1 & 0 & 8 \\ 1 & 1/2 & 0 & 1/2 & 0 & 8 \\ \hline 0 & -10 & 0 & 20 & 1 & 320 \end{array}$$

Now to make all other entries as zeros in this column, we first multiply row 1 by  $-\frac{1}{2}$  and add it to row 2, and then multiply row 1 by 10 and add it to the bottom row.

$$\begin{array}{ccccc|c} x_1 & x_2 & y_1 & y_2 & Z & C \\ 0 & 1 & 2 & -1 & 0 & 8 \\ 1 & 0 & -1 & 1 & 0 & 4 \\ \hline 0 & 0 & 20 & 10 & 1 & 400 \end{array}$$

We no longer have negative entries in the bottom row, therefore we are finished.

**Why are we finished when there are no negative entries in the bottom row?**

The answer lies in the bottom row. The bottom row corresponds to the equation:

$$0x_1 + 0x_2 + 20y_1 + 10y_2 + Z = 400 \quad \text{or} \quad Z = 400 - 20y_1 - 10y_2$$

Since all variables are non-negative, the highest value  $Z$  can ever achieve is 400, and that will happen only when  $y_1$  and  $y_2$  are zero.

8. **Read off your answers.** We now read off our answers, that is, we determine the basic solution associated with the final simplex tableau. Again, we look at the columns that have a 1 and all other entries zeros. Since the columns labeled  $y_1$  and  $y_2$  are not such columns, we arbitrarily choose  $y_1 = 0$ , and  $y_2 = 0$ , and we get

$$\left[ \begin{array}{ccc|c} x_1 & x_2 & Z & C \\ 0 & 1 & 0 & 8 \\ 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 400 \end{array} \right]$$

The matrix reads  $x_2 = 8$ ,  $x_1 = 4$ , and  $Z = 400$ .

The final solution says that if Niki works 4 hours at Job I and 8 hours at Job II, she will maximize her income to \$400. Since both slack variables are zero, it means that she would have used up all the working time, as well as the preparation time, and none will be left.

### 4.3 Minimization by the Simplex Method

In this section, you will learn to solve linear programming minimization problems using the simplex method.

1. Identify and set up a linear program in standard minimization form.
2. Formulate a dual problem in standard maximization form.
3. Use the simplex method to solve the dual maximization problem.
4. Identify the optimal solution to the original minimization problem from the optimal simplex tableau.

In this section, we will solve the standard linear programming minimization problems using the simplex method. Once again, we remind the reader that in the standard minimization problems all constraints are of the form  $ax + by \geq c$ .

The procedure to solve these problems was developed by Dr. John Von Neuman. It involves solving an associated problem called the dual problem. To every minimization problem there corresponds a dual problem. The solution of the dual problem is used to find the solution of the original problem. The dual problem is a maximization problem, which we learned to solve in the last section. We first solve the dual problem by the simplex method.

From the final simplex tableau, we then extract the solution to the original minimization problem. Before we go any further, however, we first learn to convert a minimization problem into its corresponding maximization problem called its dual.

**Example 4.3.1.** *Convert the following minimization problem into its dual.*

*Minimize*

$$Z = 12x_1 + 16x_2$$

*Subject to:*

$$x_1 + 2x_2 \geq 40$$

$$x_1 + x_2 \geq 30$$

$$x_1 \geq 0; \quad x_2 \geq 0$$

**Solution 4.3.1.** To achieve our goal, we first express our problem as the following matrix.

$$\begin{array}{cc|c} 1 & 2 & 40 \\ 1 & 1 & 30 \\ \hline 12 & 16 & 0 \end{array}$$

Observe that this table looks like an initial simplex tableau without the slack variables. Next, we write a matrix whose columns are the rows of this matrix, and the rows are the columns. Such a matrix is called a transpose of the original matrix. We get:

$$\begin{array}{cc|c} 1 & 1 & 12 \\ 2 & 1 & 16 \\ \hline 40 & 30 & 0 \end{array}$$

The following maximization problem associated with the above matrix is called its dual.

Maximize

$$Z = 40y_1 + 30y_2$$

Subject to:

$$\begin{aligned} y_1 + y_2 &\leq 12 \\ 2y_1 + y_2 &\leq 16 \\ y_1 &\geq 0; \quad y_2 \geq 0 \end{aligned}$$

Note that we have chosen the variables as  $y$ 's, instead of  $x$ 's, to distinguish the two problems.

**Example 4.3.2.** Solve both the minimization problem and its dual maximization problem graphically.

**Solution 4.3.2.** Our minimization problem is as follows.

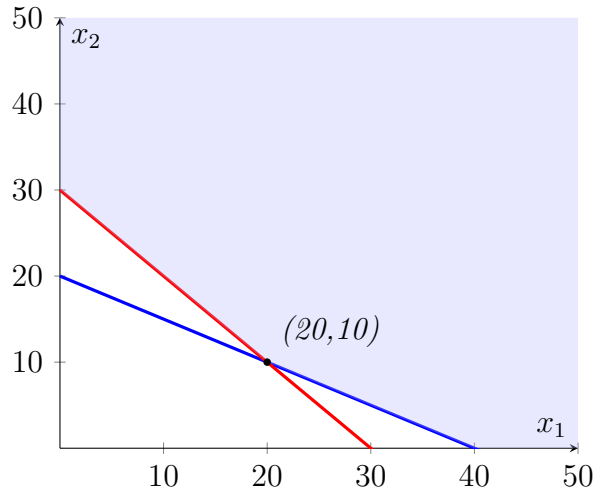
Minimize

$$Z = 12x_1 + 16x_2$$

Subject to:

$$\begin{aligned} x_1 + 2x_2 &\geq 40 \\ x_1 + x_2 &\geq 30 \\ x_1 &\geq 0; \quad x_2 \geq 0 \end{aligned}$$

We now graph the inequalities:



We have plotted the graph, shaded the feasibility region, and labeled the corner points. The corner point  $(20, 10)$  gives the lowest value for the objective function and that value is 400.

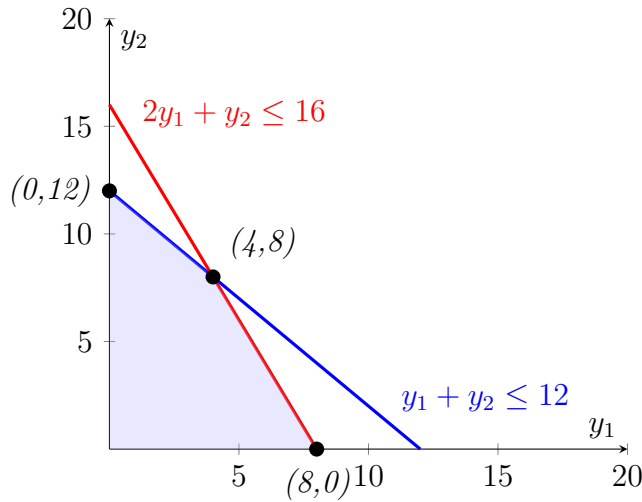
Now its dual is: Maximize

$$Z = 40y_1 + 30y_2$$

Subject to:

$$\begin{aligned} y_1 + y_2 &\leq 12 \\ 2y_1 + y_2 &\leq 16 \\ y_1 &\geq 0; \quad y_2 \geq 0 \end{aligned}$$

We graph the inequalities:



Again, we have plotted the graph, shaded the feasibility region, and labeled the corner points. The corner point  $(4, 8)$  gives the highest value for the objective function, with a value of 400.

The reader may recognize that Example 4.3.2 above is the same as Example 3.1.1 in section 3.1. It is also the same problem as Example 4.2.1 in section 4.2, where we solved it by the simplex method.

We observe that the minimum value of the minimization problem is the same as the maximum value of the maximization problem; in Example 4.3.2 the minimum and maximum are both 400. This is not a coincidence. We state the duality principle.

**Definition 4.3.1. The Duality Principle** *The objective function of the minimization problem reaches its minimum if and only if the objective function of its dual reaches its maximum. And when they do, they are equal.*

Our next goal is to extract the solution for our minimization problem in Example 4.3.1 from the corresponding dual. To do this, we solve the dual by the simplex method.

**Example 4.3.3.** *Find the solution to the minimization problem in Example 4.3.1 by solving its dual using the simplex method. We rewrite our problem:*  
Minimize

$$Z = 12x_1 + 16x_2$$

Subject to:

$$\begin{aligned}x_1 + 2x_2 &\geq 40 \\x_1 + x_2 &\geq 30 \\x_1 &\geq 0; \quad x_2 \geq 0\end{aligned}$$

**Solution 4.3.3.** *The dual is: Maximize*

$$Z = 40y_1 + 30y_2$$

Subject to:

$$\begin{aligned}y_1 + y_2 &\leq 12 \\2y_1 + y_2 &\leq 16 \\y_1 &\geq 0; \quad y_2 \geq 0\end{aligned}$$

Recall that we solved the above problem by the simplex method in Example 4.2.1, section 4.2. Therefore, we only show the initial and final simplex tableau.

The initial simplex tableau is:

$y_1$	$y_2$	$x_1$	$x_2$	$Z$	$C$	
1	1	1	0	0	12	$y_1$
2	1	0	1	0	16	$y_2$
-40	-30	0	0	1	0	$Z$

Observe an important change. Here our main variables are  $y_1$  and  $y_2$  and the slack variables are  $x_1$  and  $x_2$ .

The final simplex tableau reads as follows:

$y_1$	$y_2$	$x_1$	$x_2$	$Z$	$C$	
0	1	2	-1	0	8	
1	0	-1	1	0	4	
0	0	20	10	1	400	

A closer look at this table reveals that the  $x_1$  and  $x_2$  values along with the minimum value for the minimization problem can be obtained from the last row of the final tableau. We have highlighted these values by the arrows.



$y_1$	$y_2$	$x_1$	$x_2$	$Z$	$C$
0	1	2	-1	0	8
1	0	-1	1	0	4
<hr/>					
0	0	20	10	1	400
		↑	↑		↑

We restate the solution as follows: The minimization problem has a minimum value of 400 at the corner point  $(20, 10)$ .

#### Summary 4.3.1: Minimization by the Simplex Method

1. Set up the problem.
2. Write a matrix whose rows represent each constraint with the objective function as its bottom row.
3. Write the transpose of this matrix by interchanging the rows and columns.
4. Now write the dual problem associated with the transpose.
5. Solve the dual problem by the simplex method learned in section [4.2](#).
6. The optimal solution is found in the bottom row of the final matrix in the columns corresponding to the slack variables, and the minimum value of the objective function is the same as the maximum value of the dual.



## Chapter 5

# Exponential and Logarithmic Functions

In this chapter, you will:

1. Examine exponential and logarithmic functions and their properties.
2. Identify exponential growth and decay functions and use them to model applications.
3. Use the natural base  $e$  to represent exponential functions.
4. Use logarithmic functions to solve equations involving exponential functions.

### 5.1 Exponential Growth and Decay Models

In this section, you will learn to:

1. Recognize and model exponential growth and decay.
2. Compare linear and exponential growth.
3. Distinguish between exponential and power functions.

### 5.1.1 Comparing Exponential and Linear Growth

Consider two social media sites that are expanding the number of users they have:

- Site A has 10,000 users and expands by adding 1,500 new users each month.
- Site B has 10,000 users and expands by increasing the number of users by 10% each month.

The number of users for Site A can be modeled as linear growth. The number of users increases by a constant number, 1500, each month. If  $x$  represents the number of months that have passed and  $y$  is the number of users, the number of users after  $x$  months is given by  $y = 10000 + 1500x$ .

For Site B, the user base expands by a constant percent each month, rather than by a constant number. Growth that occurs at a constant percent each unit of time is called exponential growth.

We can compare the growth for each site by examining the number of users for the first 12 months. The table shows the calculations for the first 4 months only, but the same calculation process is used to complete the remaining months.

Month	Users at Site A	Users at Site B
0	10000	10000
1	$10000 + 1500 = 11500$	$10000 + 10\% \text{ of } 10000$ $= 10000 + 0.10(10000) = 10000(1.10) = 11000$
2	$11500 + 1500 = 13000$	$11000 + 10\% \text{ of } 11000$ $= 11000 + 0.10(11000) = 11000(1.10) = 12100$
3	$13000 + 1500 = 14500$	$12100 + 10\% \text{ of } 12100$ $= 12100 + 0.10(12100) = 12100(1.10) = 13310$
4	$14500 + 1500 = 16000$	$13310 + 10\% \text{ of } 13310$ $= 13310 + 0.10(13310) = 13310(1.10) = 14641$
5	17500	16105
6	19000	17716
7	20500	19487
8	22000	21436
9	23500	23579
10	25000	25937
11	26500	28531
12	28000	31384

For Site B, we can re-express the calculations to observe the patterns and develop a formula for the number of users after  $x$  months:

$$\text{Month 1: } y = 10000(1.1) = 11000$$

$$\text{Month 2: } y = 11000(1.1) = 10000(1.1)(1.1) = 10000(1.1)^2 = 12100$$

$$\text{Month 3: } y = 12100(1.1) = 10000(1.1)^2(1.1) = 10000(1.1)^3 = 13310$$

$$\text{Month 4: } y = 13310(1.1) = 10000(1.1)^3(1.1) = 10000(1.1)^4 = 14641$$

By observing the patterns in the calculations for months 2, 3, and 4, we can generalize the formula. After  $x$  months, the number of users  $y$  is given by the function  $y = 10000(1.1)^x$ .

### 5.1.2 Using Exponential Functions to Model Growth and Decay

In exponential growth, the value of the dependent variable  $y$  increases at a constant percentage rate as the value of the independent variable  $x$  (or  $t$ ) increases. Examples of exponential growth functions include:

- The number of residents of a city or nation that grows at a constant percent rate.
- The amount of money in a bank account that earns interest if money is deposited at a single point in time and left in the bank to compound without any withdrawals.

In exponential decay, the value of the dependent variable  $y$  decreases at a constant percentage rate as the value of the independent variable  $x$  (or  $t$ ) increases. Examples of exponential decay functions include:

- Value of a car or equipment that depreciates at a constant percent rate over time.
- The amount a drug that still remains in the body as time passes after it is ingested.
- The amount of radioactive material remaining over time as a radioactive substance decays.

Exponential functions often model quantities as a function of time; thus, we often use the letter  $t$  as the independent variable instead of  $x$ .

#### Summary 5.1.1: Exponential Growth

1. Quantity grows by a constant percent per unit of time.
2.  $y = ab^x$
3.  $a$  is a positive number representing the initial value of the function when  $x = 0$ .
4.  $b$  is a real number that is greater than 1:  $b > 1$ .
5. The growth rate  $r$  is a positive number,  $r > 0$  where  $b = 1 + r$  (so that  $r = b - 1$ ).

**Summary 5.1.2: Exponential Decay**

1. Quantity decreases by a constant percent per unit of time.
2.  $y = ab^x$
3.  $a$  is a positive number representing the initial value of the function when  $x = 0$ .
4.  $b$  is a real number that is between 0 and 1:  $0 < b < 1$ .
5. The decay rate  $r$  is a negative number,  $r < 0$  where  $b = 1 + r$  (so that  $r = b - 1$ ).

In general, the domain of exponential functions is the set of all real numbers. The range of an exponential growth or decay function is the set of all positive real numbers.

In most applications, the independent variable  $x$  or  $t$  represents time. When the independent variable represents time, we may choose to restrict the domain so that the independent variable can have only non-negative values for the application to make sense. If we restrict the domain, then the range is also restricted.

- For an exponential growth function  $y = ab^x$  with  $b > 1$  and  $a > 0$ , if we restrict the domain so that  $x \geq 0$ , then the range is  $y \geq a$ .
- For an exponential decay function  $y = ab^x$  with  $0 < b < 1$  and  $a > 0$ , if we restrict the domain so that  $x \geq 0$ , then the range is  $0 < y \leq a$ .

**Example 5.1.1.** Consider the growth models for social media sites  $A$  and  $B$ , where  $x$  = number of months since the site was started and  $y$  = number of users. The number of users for Site  $A$  follows the linear growth model:

$$y = 10000 + 1500x.$$

The number of users for Site  $B$  follows the exponential growth model:

$$y = 10000 \cdot (1.1)^x$$

For each site, use the function to calculate the number of users at the end of the first year, to verify the values in the table. Then use the functions to predict the number of users after 30 months.

**Solution 5.1.1.** Since  $x$  is measured in months, then  $x = 12$  at the end of one year.

**Linear Growth Model:** When  $x = 12$  months, then  $y = 10000 + 1500(12) = 28000$  users. When  $x = 30$  months, then  $y = 10000 + 1500(30) = 55000$  users.

**Exponential Growth Model:** When  $x = 12$  months, then  $y = 10000 \cdot (1.1)^{12} = 31384$  users. When  $x = 30$  months, then  $y = 10000 \cdot (1.1)^{30} = 174494$  users.

We see that as  $x$ , the number of months, gets larger, the exponential growth function grows larger faster than the linear function (even though in the initial stages the linear function grew faster). This is an important characteristic of exponential growth: exponential growth functions always grow faster and larger in the long run than linear growth functions.

It is helpful to use function notation, writing  $y = f(x) = ab^x$ , to specify the value of  $x$  at which the function is evaluated.

**Example 5.1.2.** A forest has a population of 2000 squirrels that is increasing at the rate of 3% per year. Let  $t$  be the number of years and  $y = f(t)$  the number of squirrels at time  $t$ .

1. Find the exponential growth function that models the number of squirrels in the forest at the end of  $t$  years.
2. Use the function to find the number of squirrels after 5 years and after 10 years.

**Solution 5.1.2.** The exponential decay function is  $y = g(t) = ab^t$ , , where the initial population  $a = 2000$  squirrels, and the growth rate  $r = 3\% = 0.03$  per year, hence  $b = 1 + r = 1.03$ .

1. The exponential growth function is  $y = f(t) = 2000(1.03^t)$
2. After 5 years, the population is  $y = f(5) = 2000(1.03^5) \approx 2315.25$  squirrels. After 10 years, the population is  $y = f(10) = 2000(1.03^{10}) \approx 2691.70$  squirrels.

**Example 5.1.3.** A large lake has a population of 1000 frogs. Unfortunately, the frog population is decreasing at the rate of 5% per year. Let  $t$  represent the number of years and  $y = g(t)$  the number of frogs in the lake at time  $t$ .

1. Find the exponential decay function that models the frog population.
2. Calculate the population size after 10 years.



**Solution 5.1.3.** The exponential decay function is  $y = g(t) = ab^t$ , where  $a = 1000$  is the initial population of frogs, and the decay rate is 5% per year, which translates to  $r = -0.05$ . Therefore,  $b = 1 + r = 0.95$ .

1. The function modeling the frog population is  $y = g(t) = 1000(0.95)^t$ .
2. After 10 years, the population is  $y = g(10) = 1000(0.95)^{10} \approx 599$  frogs, showing a significant decrease due to the yearly decline.

**Example 5.1.4.** A bacteria population is described by the function  $y = f(t) = 100(2^t)$ , where  $t$  is the time in hours, and  $y$  is the number of bacteria.

1. What is the initial population?
2. What is change after the first hour?
3. How long does it take for the population to reach 800 bacteria?

**Solution 5.1.4.**

1. The initial population is  $y = f(0) = 100(2^0) = 100$  bacteria, as  $a = 100$  in the function  $f(t)$ .
2. After one hour, the population doubles to  $y = f(1) = 100(2^1) = 200$  bacteria.
3. To find when the population reaches 800, solve  $800 = 100(2^t)$ :

$$800 = 100(2^t)$$

$$8 = 2^t$$

$$2^3 = 2^t$$

$$\Rightarrow t = 3.$$

*It takes 3 hours for the population to grow to 800 bacteria.*

Important notes about Example 5.1.4:

1. In solving  $8 = 2t$ , we "knew" that  $t$  is 3. But we usually cannot know the value of the variable just by looking at the equation. Later, we will use logarithms to solve equations that have the variable in the exponent.

2. To solve  $800 = 100(2^t)$ , we divided both sides by 100 to isolate the exponential expression  $2^t$ . We cannot multiply 100 by 2. Even if we write it as  $800 = 100(2)t$ , which is equivalent, we still cannot multiply 100 by 2. The exponent applies only to the quantity immediately before it, so the exponent  $t$  applies only to the base of 2.

### 5.1.3 Comparing Linear, Exponential, and Power Functions

To identify the type of function from its formula, we need to carefully note the position that the variable occupies in the formula.

#### Summary 5.1.3: Comparing Functions

- A **linear function** can be written in the form  $y = ax + b$ . As we studied in Chapter 1, there are other forms in which linear equations can be written, but linear functions can all be rearranged to have the form  $y = mx + b$ .
- An **exponential function** has the form  $y = ab^x$ , where the variable  $x$  is in the exponent. The base  $b$  is a positive number:
  - If  $b > 1$ , the function represents exponential growth.
  - If  $0 < b < 1$ , the function represents exponential decay.
- A **power function** has the form  $y = cx^p$ , where the variable  $x$  is in the base. The exponent  $p$  is a non-zero number.

We compare three functions:

- linear function  $y = f(x) = 2x$
- exponential function  $y = g(x) = 2^x$
- power function  $y = h(x) = x^2$

$x$	$f(x) = 2x$	$g(x) = 2^x$	$h(x) = x^2$
0	0	1	0
1	2	2	1
2	4	4	4
3	6	8	9
4	8	16	16
5	10	32	25
6	12	64	36
10	20	1024	100

For the functions in the previous table: the linear function  $y = f(x) = 2x$ , the exponential function  $y = g(x) = 2^x$ , and the power function  $y = h(x) = x^2$ , if we restrict the domain to  $x \geq 0$  only, then all these functions are growth functions. When  $x \geq 0$ , the value of  $y$  increases as the value of  $x$  increases.

The exponential growth function grows larger faster than the linear and power functions as  $x$  gets large. This is always true of exponential growth functions as  $x$  gets large enough.

Notice that for equal intervals of change in  $x$

- with a linear function  $y$  increases by addition of a constant amount.
- with an exponential function,  $y$  is multiplied by a constant amount.
- with a power function neither of these is true.

**Example 5.1.5.** *Classify the functions below as exponential, linear, or power functions.*

**a.**  $y = 10x^3$

**b.**  $y = 1000 - 30x$

**c.**  $y = 1000(1.05^x)$

**d.**  $y = 500(0.75^x)$

**e.**  $y = 10\sqrt[3]{x} = x^{1/3}$

**f.**  $y = 5x - 1$

**g.**  $y = \frac{6}{x^2} = 6x^{-2}$

**Solution 5.1.5.** *The exponential functions are*

**c.**  $y = 1000(1.05^x)$  The variable is in the exponent; the base is the number  $b = 1.05$ .

**d.**  $y = 500(0.75^x)$  The variable is in the exponent; the base is the number  $b = 0.75$ .

The linear functions are

**b.**  $y = 1000 - 30x$

**f.**  $y = 5x - 1$

The power functions are

**a.**  $y = 10x^3$  The variable is the base; the exponent is a fixed number,  $p = 3$ .

**e.**  $y = 10\sqrt[3]{x} = 10x^{1/3}$  The variable is the base; the exponent is a number,  $p = 1/3$ .

**g.**  $y = \frac{6}{x^2} = 6x^{-2}$  The variable is the base; the exponent is a number,  $p = -2$ .

### 5.1.4 The Natural Base

The number  $e$  is often used as the base of an exponential function and is called the natural base. It is approximately 2.71828 and is an irrational number with an infinite never repeating decimal expansion. The reader may be familiar with another famous irrational number,  $\pi$ . Section 6.2.4 shows how the value of  $e$  enters the world of finance and why this number is mathematically important.

When  $e$  is the base in an exponential growth or decay function, it is referred to as continuous growth or continuous decay. We will use  $e$  in Chapter 6 in financial calculations when we examine interest that compounds continuously.

Any exponential function can be written in the form  $y = ae^{kx}$ , where  $k$  is called the continuous growth or decay rate:

- If  $k > 0$ , the function represents exponential growth.
- If  $k < 0$ , the function represents exponential decay.

The initial value  $a$  is the starting amount of whatever is growing or decaying.

We can rewrite the function in the form  $y = ab^x$ , where  $b = e^k$ .

In general, if we know one form of the equation, we can find the other forms. For now, we have not yet covered the skills to find  $k$  when we know  $b$ . After we learn about logarithms later in this chapter, we will find  $k$  using the natural logarithm:  $k = \ln(b)$ .

The table below summarizes the forms of exponential growth and decay functions.

	$y = ab^x$	$y = a(1 + r)^x$	$y = ae^{kx}, k \neq 0$
Initial value	$a > 0$	$a > 0$	$a > 0$
Relationship between $b, r, k$	$b > 0$	$b = 1 + r$	$b = e^k$ and $k = \ln b$
Growth	$b > 1$	$r > 0$	$k > 0$
Decay	$0 < b < 1$	$r < 0$	$k < 0$

**Example 5.1.6.** *The value of houses in a city are increasing at a continuous growth rate of 6% per year. For a house that currently costs \$400,000:*

1. *Write the exponential growth function in the form  $y = ae^{kx}$ .*
2. *What would be the value of this house 4 years from now?*
3. *Rewrite the exponential growth function in the form  $y = ab^x$ .*
4. *Find and interpret  $r$ .*

**Solution 5.1.6.**

1. *The initial value of the house is  $a = \$400,000$ . The problem states that the continuous growth rate is 6% per year, so  $k = 0.06$ . The growth function is:  $y = 400000e^{0.06x}$ .*
2. *After 4 years, the value of the house is  $y = 400000e^{0.06(4)} \approx \$508,500$ .*
3. *To rewrite  $y = 400000e^{0.06x}$  in the form  $y = ab^x$ , we use the fact that  $b = e^k$ . Therefore,  $b = e^{0.06} \approx 1.061836547 \approx 1.0618$  and the function becomes  $y = 400000(1.0618)^x$ .*
4. *To find  $r$ , we use the fact that  $b = 1 + r$ . Given  $b = 1.0618$ , we solve  $1 + r = 1.0618$  which gives  $r = 0.0618$ . The value of the house is*

increasing at an annual rate of 6.18%.

**Example 5.1.7.** Suppose that the value of a certain model of new car decreases at a continuous decay rate of 8% per year. For a car that costs \$20,000 when new:

1. Write the exponential decay function in the form  $y = ae^{kx}$ .
2. What would be the value of this car 5 years from now?
3. Rewrite the exponential decay function in the form  $y = ab^x$ .
4. Find and interpret  $r$ .

**Solution 5.1.7.** 1. The initial value of the car is  $a = \$20,000$ . The problem states that the continuous decay rate is 8% per year, so  $k = -0.08$ . The decay function is:  $y = 20000e^{-0.08x}$ .

2. After 5 years, the value of the car is  $y = 20000e^{-0.08(5)} \approx \$13,406.40$ .
3. To rewrite  $y = 20000e^{-0.08x}$  in the form  $y = ab^x$ , we use the fact that  $b = e^k$ . Therefore,  $b = e^{-0.08} \approx 0.9231163464 \approx 0.9231$  and the function becomes  $y = 20000(0.9231)^x$ .
4. To find  $r$ , we use the fact that  $b = 1 + r$ . Given  $b = 0.9231$ , we solve  $1 + r = 0.9231$  which gives  $r = -0.0769$ . The value of the car is decreasing at an annual rate of 7.69%.

## 5.2 Graphing Exponential Functions

In this section, you will:

1. examine properties of exponential functions
2. examine graphs of exponential functions

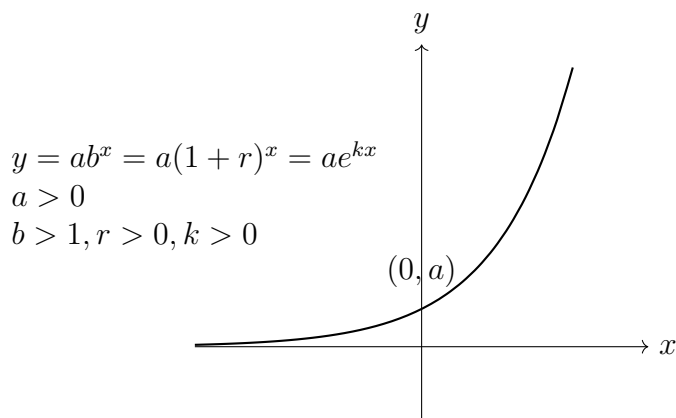
An exponential function can be written in forms  $f(x) = ab^x = a(1 + r)^x = ae^{kx}$ :

- $a$  is the initial value because  $f(0) = a$ .
- In the growth and decay models that we examine in this finite math textbook,  $a > 0$ .

- $b$  is often called the growth factor. We restrict  $b$  to be positive ( $b > 0$ ) because even roots of negative numbers are undefined. We want the function to be defined for all values of  $x$ , but  $b^x$  would be undefined for some values of  $x$  if  $b < 0$ .
- $r$  is called the growth or decay rate. In the formula for the functions, we use  $r$  in decimal form, but in the context of a problem we usually state  $r$  as a percent.
- $k$  is called the continuous growth rate or continuous decay rate.

### 5.2.1 Properties of Exponential Growth Functions

- The function  $y = f(x) = ab^x$  represents growth if  $b > 1$  and  $a > 0$ .
- The growth rate  $r$  is positive when  $b > 1$ . Because  $b = 1 + r > 1$ , then  $r = b - 1 > 0$ .
- The function  $y = f(x) = ae^{kx}$  represents growth if  $k > 0$  and  $a > 0$ .
- The function is an increasing function;  $y$  increases as  $x$  increases.



There are some properties the reader should notice:

- **Domain:** All real numbers can be input to an exponential function. Mathematicians would write  $\{x \in \mathbb{R}\}$  or simply  $\mathbb{R}$ .
- **Range:** If  $a > 0$ , the range is the set of all positive real numbers. Mathematicians would write  $\{y \in \mathbb{R} : y > 0\}$ .
- The graph is always above the  $x$  axis.

- **Horizontal Asymptote:** when  $b > 1$ , the horizontal asymptote is the negative  $x$  axis, as  $x$  becomes large negative. Using mathematical notation: as  $x \rightarrow -\infty$ , then  $y \rightarrow 0$ .
- The vertical intercept is the point  $(0, a)$  on the  $y$ -axis.
- There is no horizontal intercept because the function does not cross the  $x$ -axis.

### 5.2.2 Properties of Exponential Decay Functions

- The function  $y = f(x) = ab^x$  represents decay if  $0 < b < 1$  and  $a > 0$ .
- The growth rate  $r$  is negative when  $0 < b < 1$ . Because  $b = 1 + r < 1$ , then  $r = b - 1 < 0$ .
- The function  $y = f(x) = ae^{kx}$  represents decay if  $k < 0$  and  $a > 0$ .
- The function is a decreasing function;  $y$  decreases as  $x$  increases.
- **Domain:** All real numbers can be input to an exponential function. Mathematicians would write  $\{x \in \mathbb{R}\}$  or simply  $\mathbb{R}$ .
- **Range:** If  $a > 0$ , the range is the set of all positive real numbers. Mathematicians would write  $\{y \in \mathbb{R} : y > 0\}$ .
- **Horizontal Asymptote:** when  $b < 1$ , the horizontal asymptote is the positive  $x$  axis as  $x$  becomes large positive. Using mathematical notation: as  $x \rightarrow \infty$ , then  $y \rightarrow 0$ .
- The vertical intercept is the point  $(0, a)$  on the  $y$ -axis. There is no horizontal intercept because the function does not cross the  $x$ -axis.

The graphs for exponential growth and decay functions are displayed in the figure 5.2.2 below for comparison.

### 5.2.3 An Exponential Function Is a One-to-One Function

Observe that in the graph of an exponential function, each  $y$  value on the graph occurs only once. Therefore, every  $y$  value in the range corresponds to only one  $x$  value. So, for any particular value of  $y$ , you can use the graph



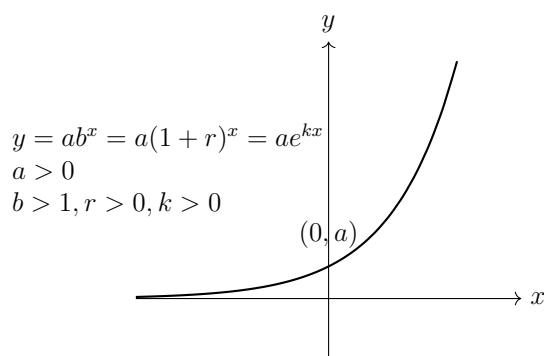


Figure 5.1: Exponential Growth

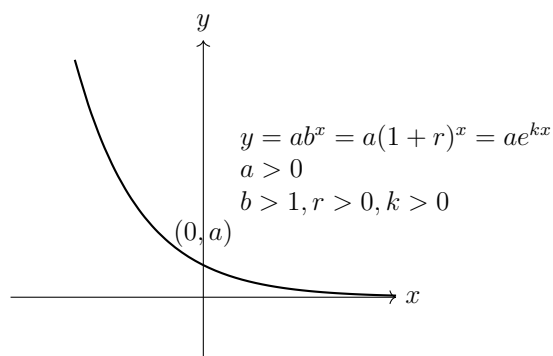


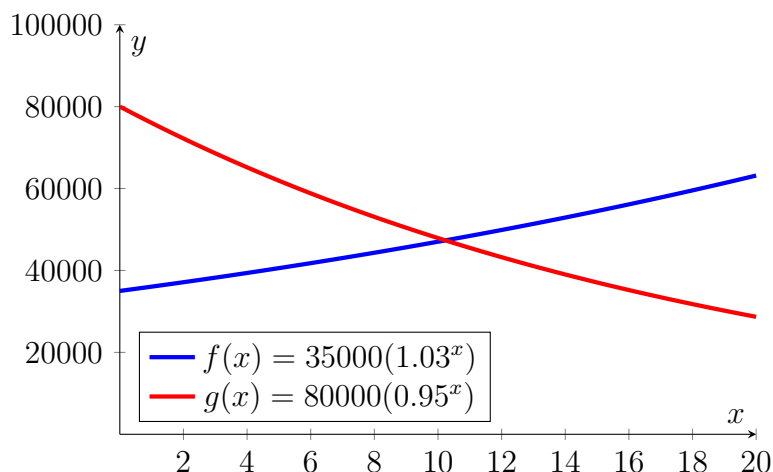
Figure 5.2: Exponential Decay

to see which value of  $x$  is the input to produce that  $y$  value as output. This property is called "one-to-one".

Because for each value of the output  $y$ , you can uniquely determine the value of the corresponding input  $x$ , thus every exponential function has an inverse function. The inverse function of an exponential function is a logarithmic function, which we will investigate in the next section.

**Example 5.2.1.**  $x$  years after the year 2025, the population of the city of Fulton is given by the function  $y = f(x) = 35000(1.03^x)$ , and  $x$  years after the year 2025, the population of the city of Greenville is given by the function  $y = g(x) = 80000(0.95^x)$ . Compare the graphs of these functions.

**Solution 5.2.1.** Here are graphs:



Some observations we may make:

- *Fulton's population is undergoing Exponential Growth.*
  - $b = 1.03 > 1$  and  $r = 0.03 > 0$
  - *The population is increasing.*
- *Greenville's population is undergoing Exponential Decay.*
  - $b = 0.95 < 1$  and  $r = -0.05 < 0$
  - *The population is decreasing.*
- *The initial population of Fulton in 2025 is 35000.*
- *The initial population of Greenville in 2025 is 80000.*
- *In general, the domains of both functions  $f(x)$  and  $g(x)$  are all real numbers, but realistically the populations of Fulton and Greenville cannot follow this model for all such values of  $x$ . Depending on how they were created, the models may be good enough for use for  $x$  representing several years in the past to several years in the future.*
- *If the model holds, the populations of the two cities will be approximately equal sometime around 2035.*

## 5.3 Logarithms and Logarithmic Functions

In this section, you will learn:

- The definition of a logarithmic function as the inverse of the exponential function.
- How to write equivalent logarithmic and exponential expressions.
- The definition of common logarithms and natural logarithms.
- Properties of logarithms and the Log Rules.

### 5.3.1 Define the Logarithm

Suppose that a population of 50 flies is expected to double every week, leading to a function of the form  $f(x) = 50 \cdot (2)^x$ , where  $x$  represents the number of weeks that have passed. When will this population reach 500? Trying to solve this problem leads to  $500 = 50 \cdot (2)^x$ . Dividing both sides by 50 to isolate the exponential leads to  $10 = 2^x$ .

While we have set up exponential models and used them to make predictions, you may have noticed that solving exponential equations has not yet been mentioned. The reason is simple: none of the algebraic tools discussed so far are sufficient to solve exponential equations. Consider the equation  $2^x = 10$  above. We know that  $2^3 = 8$  and  $2^4 = 16$ , so it is clear that  $x$  must be some value between 3 and 4 since  $g(x) = 2^x$  is increasing. We could use technology to create a table of values or graph to better estimate the solution, but we would like to find an algebraic way to solve the equation.

We need an inverse operation to exponentiation in order to solve for the variable if the variable is in the exponent. As we learned in algebra class, the inverse function for an exponential function is a logarithmic function. We also learned that an exponential function has an inverse function, because each output ( $y$ ) value corresponds to only one input ( $x$ ) value. The name given to this property was “one-to-one”.

**Definition 5.3.1.** The **logarithm** (base  $b$ ), written  $\log_b(x)$ , is the inverse of the exponential function (base  $b$ ),  $b^x$ .

$$y = \log_b(x) \quad \Leftrightarrow \quad b^y = x$$

**Some Notes**

- In general,  $b^a = c$  is called an **exponential equation** and is equivalent to the **logarithmic equation**  $\log_b(c) = a$ .
- The base  $b$  must be positive:  $b > 0$
- Since the logarithm and exponential are inverses, it follows that:

$$\log_b(b^x) = x \quad \text{and} \quad b^{\log_b(x)} = x$$

- Since log is a function, it is most correctly written as  $\log_b(c)$ , using parentheses to denote function evaluation, just as we would with  $f(c)$ . However, when the input is a single variable or number, it is common to see the parentheses dropped and the expression written as  $\log_b c$ .

**Example 5.3.1.** *Write these exponential equations as logarithmic equations:*

1.  $2^3 = 8$
2.  $5^2 = 25$
3.  $10^{-3} = \frac{1}{1000}$

**Solution 5.3.1.**

1.  $2^3 = 8$  can be written as a logarithmic equation as  $\log_2(8) = 3$
2.  $5^2 = 25$  can be written as a logarithmic equation as  $\log_5(25) = 2$
3.  $10^{-3} = \frac{1}{1000}$  can be written as a logarithmic equation as  $\log_{10}\left(\frac{1}{1000}\right) = -3$

**Example 5.3.2.** *Write these logarithmic equations as exponential equations:*

1.  $\log_6(\sqrt{6}) = \frac{1}{2}$
2.  $\log_3(9) = 2$

**Solution 5.3.2.**

1.  $\log_6(\sqrt{6}) = \frac{1}{2}$  can be written as an exponential equation as  $6^{\frac{1}{2}} = \sqrt{6}$
2.  $\log_3(9) = 2$  can be written as an exponential equation as  $3^2 = 9$

By establishing the relationship between exponential and logarithmic functions, we can now solve basic logarithmic and exponential equations by rewriting.

**Example 5.3.3.** Solve  $\log_4(x) = 2$  for  $x$ .

**Solution 5.3.3.** By rewriting this expression as an exponential,  $4^2 = x$ , so  $x = 16$ .

**Example 5.3.4.** Solve  $2^x = 10$  for  $x$ .

**Solution 5.3.4.** By rewriting this expression as a logarithm, we get  $x = \log_2(10)$ . Using a computer utility we find  $x \approx 3.32192809489$ .

While this does define a solution, you may find it somewhat unsatisfying since it is difficult to compare this expression to the decimal estimate we made earlier. Also, giving an exact expression for a solution is not always useful—often we really need a decimal approximation to the solution. Luckily, this is a task that calculators and computers are quite adept at. Unfortunately for us, most calculators will only evaluate logarithms of two bases: base 10 and base  $e$ . Computer utilities such as [Desmos](#) and [Wolfram Alpha](#) compute such things nicely. Please note, even with many decimal places computers and calculators are providing estimates.

### 5.3.2 Natural and Base 10 Logarithms

**Definition 5.3.2.** The **natural** logarithm is the logarithm with  $e$  as the base ( $\log_e(x)$ ). It is often written as  $\ln(x)$ .

**Definition 5.3.3.** The **base 10** (or common) logarithm is the logarithm with 10 as the base ( $\log_{10}(x)$ ). It is often written as  $\log(x)$ .

**Example 5.3.5.** Evaluate  $\log(1000)$  using the definition of the common log.

**Solution 5.3.5.** The table shows values of the common log:

<i>number</i>	<i>number as exponential</i>	$\log(\textit{number})$
1000	$10^3$	3
100	$10^2$	2
10	$10^1$	1
1	$10^0$	0
0.1	$10^{-1}$	-1
0.01	$10^{-2}$	-2
0.001	$10^{-3}$	-3

To evaluate  $\log(1000)$ , we can say

$$x = \log(1000)$$

Then rewrite the equation in exponential form using the common log base of 10

$$10^x = 1000$$

From this, we might recognize that 1000 is the cube of 10, so

$$x = 3.$$

Alternatively, we can use the inverse property of logs to write

$$\log_{10}(10^3) = 3$$

**Example 5.3.6.** Evaluate  $\log(1/1,000,000)$

**Solution 5.3.6.** To evaluate  $\log(1/1,000,000)$ , we can say

$$x = \log(1/1,000,000) = \log(1/10^6) = \log(10^{-6})$$

Then rewrite the equation in exponential form:

$$10^x = 10^{-6}$$

Therefore  $x = -6$ .

Alternatively, we can use the inverse property of logs to find the answer:

$$\log_{10}(10^{-6}) = -6$$

**Example 5.3.7.** Evaluate

1.  $\ln e^5$
2.  $\ln \sqrt{e}$ .

**Solution 5.3.7.**

1. To evaluate  $\ln e^5$ , we can say

$$x = \ln e^5$$

Then rewrite into exponential form using the natural log base of  $e$

$$e^x = e^5$$

Therefore  $x = 5$

Alternatively, we can use the inverse property of logs to write  $\ln(e^5) = 5$

2. To evaluate  $\ln \sqrt{e}$ , we recall that roots are represented by fractional exponents

$$x = \ln \sqrt{e} = \ln(e^{1/2})$$

Then rewrite into exponential form using the natural log base of  $e$

$$e^x = e^{1/2}$$

Therefore  $x = 1/2$

Alternatively, we can use the inverse property of logs to write  $\ln(e^{1/2}) = 1/2$

**Example 5.3.8.** Evaluate the following using your calculator or computer:

1.  $\log 500$
2.  $\ln 500$

**Solution 5.3.8.**

1.  $\log 500 \approx 2.69897$

2.  $\ln 500 \approx 6.214608$

**5.3.3 Properties of Logarithms****Summary 5.3.1:****1. Exponent Property:**

$$\log_b(A^p) = p \log_b(A)$$

**2. Product Property:**

$$\log_b(AC) = \log_b(A) + \log_b(C)$$

**3. Quotient Property:**

$$\log_b\left(\frac{A}{C}\right) = \log_b(A) - \log_b(C)$$

**5.4 Graphs and Properties of Logarithmic Functions**

In this section, you will:

1. examine properties of logarithmic functions
2. examine graphs of logarithmic functions
3. examine the relationship between graphs of exponential and logarithmic functions

Recall that the exponential function  $f(x) = 2^x$  produces this table of values

$x$	-3	-2	-1	0	1	2	3
$f(x)$	1/8	1/4	1/2	1	2	4	8

Since the logarithmic function is an inverse of the exponential,  $g(x) = \log_2(x)$



#### 5.4. GRAPHS AND PROPERTIES OF LOGARITHMIC FUNCTIONS 137

produces the table of values

$x$	$1/8$	$1/4$	$1/2$	$1$	$2$	$4$	$8$
$g(x)$	$-3$	$-2$	$-1$	$0$	$1$	$2$	$3$

In this second table, notice that:

- As the input increases, the output increases.
- As input increases, the output increases more slowly.
- Since the exponential function only outputs positive values, the logarithm can only accept positive values as inputs, so the domain of the log function is  $(0, \infty)$ .
- Since the exponential function can accept all real numbers as inputs, the logarithm can have any real number as output, so the range is all real numbers or  $(-\infty, \infty)$ .

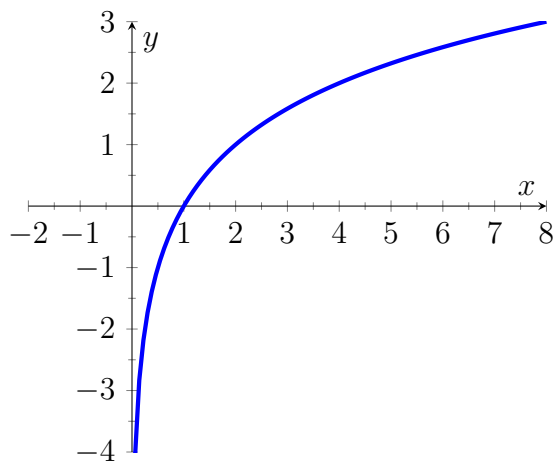
Plotting the graph of  $g(x) = \log_2(x)$  from the points in the table, notice that as the input values for  $x$  approach zero, the output of the function grows very large in the negative direction, indicating a vertical asymptote at  $x = 0$ .

In symbolic notation we write

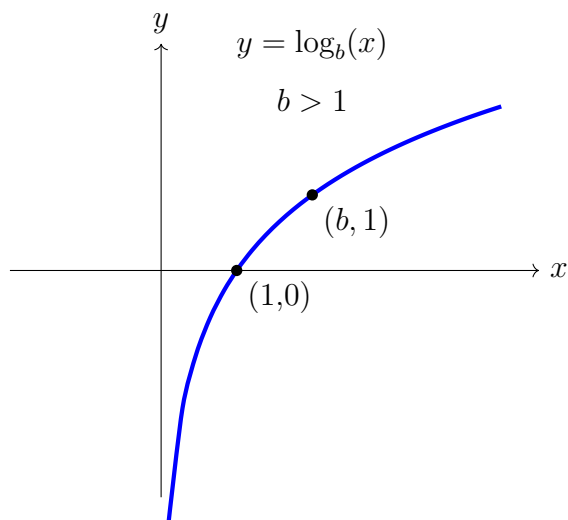
$$\text{as } x \rightarrow 0^+, f(x) \rightarrow -\infty$$

and

$$\text{as } x \rightarrow \infty, f(x) \rightarrow \infty$$

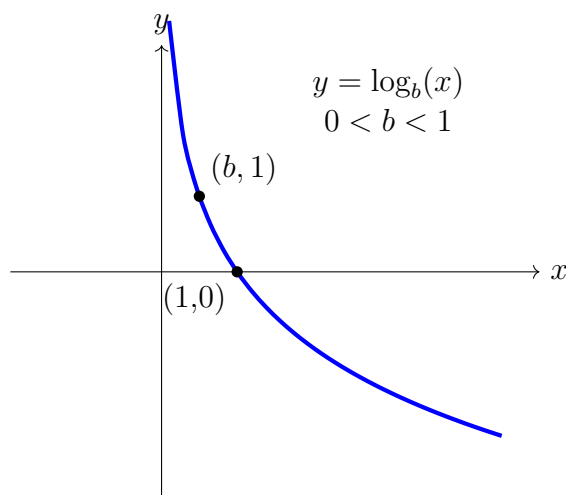


Graphically, in the function  $g(x) = \log_b(x)$ ,  $b > 1$ , we observe the following properties:



- The graph has a horizontal intercept at  $(1, 0)$ .
- The line  $x = 0$  (the y-axis) is a vertical asymptote; as  $x \rightarrow 0^+$ ,  $y \rightarrow -\infty$ .
- The graph is increasing if  $b > 1$ .
- The domain of the function is  $x > 0$ , or  $(0, \infty)$ .
- The range of the function is all real numbers, or  $(-\infty, \infty)$ .

However if the base  $b$  is less than 1,  $0 < b < 1$ , then the graph appears as below.



- The graph has a horizontal intercept at  $(1, 0)$ .
- The line  $x = 0$  (the y-axis) is a vertical asymptote; as  $x \rightarrow 0^+$ ,  $y \rightarrow -\infty$ .
- The graph is decreasing if  $0 < b < 1$ .
- The domain of the function is  $x > 0$ , or  $(0, \infty)$ .
- The range of the function is all real numbers, or  $(-\infty, \infty)$ .

When graphing a logarithmic function, it can be helpful to remember that the graph will pass through the points  $(1, 0)$  and  $(b, 1)$ .

Finally, we compare the graphs of  $y = b^x$  and  $y = \log_b(x)$ , shown below on the same axes.

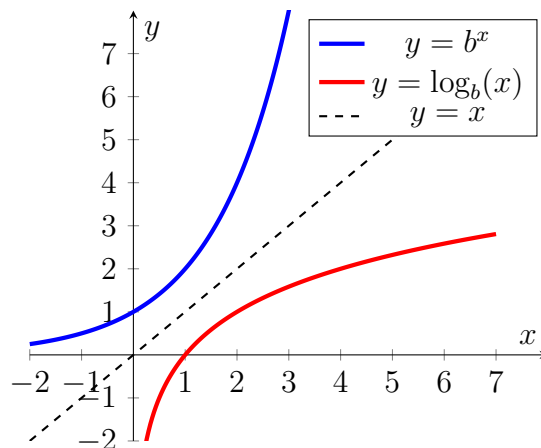
Because the functions are inverse functions of each other, for every specific ordered pair  $(h, k)$  on the graph of  $y = bx$ , we find the point  $(k, h)$  with the coordinates reversed on the graph of  $y = \log_b(x)$ .

In other words, if the point with  $x = h$  and  $y = k$  is on the graph of  $y = bx$ , then the point with  $x = k$  and  $y = h$  lies on the graph of  $y = \log_b(x)$ .

The domain of  $y = bx$  is the range of  $y = \log_b(x)$ . The range of  $y = bx$  is the domain of  $y = \log_b(x)$ .

For this reason, the graphs appear as reflections, or mirror images, of each other across the diagonal line  $y = x$ . This is because the inputs and outputs

are swapped for a function and its inverse.



## 5.5 Application Problems with Exponential and Logarithmic Functions

In this section, you will:

1. Review strategies for solving equations arising from exponential formulas.
2. Solve application problems involving exponential functions and logarithmic functions.

### 5.5.1 Strategies for Solving Equations That Contain Exponents

When solving application problems that involve exponential and logarithmic functions, we need to pay close attention to the position of the variable in the equation to determine the proper way solve the equation we investigate solving equations that contain exponents.

Suppose we have an equation in the form:

$$\text{value} = \text{coefficient}(\text{base})^{\text{exponent}}$$

We consider four strategies for solving the equation:

- If the coefficient, base, and exponent are all known, we only need to evaluate the expression for coefficient(base)<sup>exponent</sup> to evaluate its value.
- If the variable is the coefficient, evaluate the expression for (base)<sup>exponent</sup>. Then it becomes a linear equation which we solve by dividing to isolate the variable.
- If the variable is in the exponent, use logarithms to solve the equation.
- If the variable is not in the exponent, but is in the base, use roots to solve the equation.

**Example 5.5.1.** Suppose that a stock's price is rising at the rate of 7% per year, and that it continues to increase at this rate. If the value of one share of this stock is \$43 now, find the value of one share of this stock three years from now.

**Solution 5.5.1.** Let  $y$  be the value of the stock after  $t$  years:  $y = ab^t$ .

The problem tells us that  $a = 43$  and  $r = 0.07$ , so  $b = 1 + r = 1 + 0.07 = 1.07$ .

Therefore, function  $y$  is  $43(1.07)^t$ .

In this case we know that  $t = 3$  years, and we need to evaluate  $y$  when  $t = 3$ .

At the end of 3 years, the value of this one share of this stock will be

$$y = 43(1.07)^3 = \$52.68$$

**Example 5.5.2.** The value of a new car depreciates (decreases) after it is purchased. Suppose that the value of the car depreciates according to an exponential decay model. Suppose that the value of the car is \$12000 at the end of 5 years and that its value has been decreasing at the rate of 9% per year. Find the value of the car when it was new.

**Solution 5.5.2.** Let  $y$  be the value of the car after  $t$  years:  $y = ab^t$ .

Given  $r = -0.09$  and  $b = 1 + r = 1 + (-0.09) = 0.91$ .

The function is  $y = a(0.91)^t$ .

For  $t = 5$ , and  $y = 12000$ ; substituting these values gives  $12000 = a(0.91)^5$ .

Solve for  $a$ :

First, evaluate  $(0.91)^5$  and then solve the resulting linear equation to find  $a$ .

$$12000 = a(.624)$$

$$a = \frac{12000}{0.624} = \$19,230.77$$

The car's value was \$19,230.77 when it was new.

**Example 5.5.3.** A national park has a population of 5000 deer in the year 2024. Conservationists are concerned because the deer population is decreasing at the rate of 7% per year. If the population continues to decrease at this rate, how long will it take until the population is only 3000 deer?

**Solution 5.5.3.** Let  $y$  be the number of deer in the national park  $t$  years after the year 2024:  $y = ab^t$ .

Given  $r = -0.07$  and  $b = 1 + r = 1 + (-0.07) = 0.93$ , the initial population is  $a = 5000$ .

The exponential decay function is  $y = 5000(0.93)^t$ .

To find when the population will be 3000, substitute  $y = 3000$ :

$$3000 = 5000(0.93)^t$$

Next, divide both sides by 5000 to isolate the exponential expression:

$$\frac{3000}{5000} = (0.93)^t$$

$$0.6 = 0.93^t$$

Rewrite the equation in logarithmic form; then use the change of base formula to evaluate.

$$t = \log_{0.93}(0.6)$$

$$t = \frac{\ln(0.6)}{\ln(0.93)} \approx 7.039 \text{ years}$$

After 7.039 years, there are 3000 deer.

In Example 5.5.3, we needed to state the answer to several decimal places of precision to remain accurate. Evaluating the original function using a rounded value of  $t = 7$  years gives a value that is close to 3000, but not exactly 3000.

$$y = 5000(0.93)^7 = 3008.5 \text{ deer}$$

However using  $t = 7.039$  years produces a value of 3000 for the population of deer

$$y = 5000(0.93)^{7.039} = 3000.0016 \approx 3000 \text{ deer}$$

**Example 5.5.4.** A video posted on YouTube initially had 80 views as soon as it was posted. The total number of views to date has been increasing exponentially according to the exponential growth function  $y = 80e^{0.2t}$ , where  $t$  represents time measured in days since the video was posted. How many days does it take until 2500 people have viewed this video?

**Solution 5.5.4.** Let  $y$  be the total number of views  $t$  days after the video is initially posted. We are given that the exponential growth function is  $y = 80e^{0.2t}$  and we want to find the value of  $t$  for which  $y = 2500$ . Substitute  $y = 2500$  into the equation and use natural log to solve for  $t$ .

First, divide both sides of the equation by 80 to isolate the exponential term:

$$\frac{2500}{80} = e^{0.2t}$$

This simplifies to:

$$31.25 = e^{0.2t}$$

Taking the natural logarithm of both sides gives us:

$$0.2t = \ln(31.25)$$

Dividing both sides by 0.2 to solve for  $t$  yields:

$$t = \frac{\ln(31.25)}{0.2}$$

Using a calculator, we find that:

$$t \approx \frac{3.442}{0.2} \approx 17.2 \text{ days}$$

Thus, the video will have 2500 total views approximately 17 days after it was posted.

**Example 5.5.5.** A statistician creates a website to analyze sports statistics. His business plan states that his goal is to accumulate 50,000 followers by the end of 2 years (24 months from now). He hopes that if he achieves this goal his site will be purchased by a sports news outlet. The initial user base of people signed up as a result of pre-launch advertising is 400 people. Find the monthly growth rate needed if the user base is to accumulate to 50,000 users at the end of 24 months.

**Solution 5.5.5.** Let  $y$  be the total user base  $t$  months after the site is launched.

The growth function for this site is  $y = 400(1 + r)^t$ .

We don't know the growth rate  $r$ . We do know that when  $t = 24$  months, then  $y = 50000$ .

Substitute the values of  $y$  and  $t$ ; then we need to solve for  $r$ .

$$\frac{50000}{400} = (1 + r)^{24}$$

Divide both sides by 400 to isolate  $(1 + r)^{24}$  on one side of the equation

$$125 = (1 + r)^{24}$$

Because the variable in this equation is in the base, we use roots:

$$\sqrt[24]{125} = 1 + r$$



$$125^{1/24} = 1 + r$$

$$1.2228 \approx 1 + r$$

$$r \approx 0.2228$$

The website's user base needs to increase at the rate of 22.28% per month in order to accumulate 50,000 users by the end of 24 months.

**Example 5.5.6.** A fact sheet on caffeine dependence from Johns Hopkins Medical Center states that the half-life of caffeine in the body is between 4 and 6 hours. Assuming that the typical half-life of caffeine in the body is 5 hours for the average person and that a typical cup of coffee has 120 mg of caffeine.

1. Write the decay function.
2. Find the hourly rate at which caffeine leaves the body.
3. How long does it take until only 20 mg of caffeine is still in the body?

**Solution 5.5.6.**

1. Let  $y$  be the total amount of caffeine in the body  $t$  hours after drinking the coffee. Exponential decay function  $y = ab^t$  models this situation. The initial amount of caffeine is  $a = 120$ .

We don't know  $b$  or  $t$ , but we know that the half-life of caffeine in the body is 5 hours. This tells us that when  $t = 5$ , then there is half the initial amount of caffeine remaining in the body.

$$y = 120b^t$$

$$\frac{1}{2}(120) = 120b^5$$

$$60 = 120b^5$$

Divide both sides by 120 to isolate the expression  $b^5$  that contains the variable.

$$\frac{60}{120} = b^5$$

$$0.5 = b^5$$

The variable is in the base and the exponent is a number. Use roots to solve for  $b$ :

$$b = \sqrt[5]{0.5}$$

$$b \approx 0.87$$

We can now write the decay function for the amount of caffeine (in mg.) remaining in the body  $t$  hours after drinking a cup of coffee with 120 mg of caffeine

$$y = f(t) = 120(0.87)^t$$

2. Use  $b = 1 + r$  to find the decay rate  $r$ . Because  $b = 0.87$  and the amount of caffeine in the body is decreasing over time, the value of  $r$  will be negative.

$$0.87 = 1 + r$$

$$r = -0.13$$

The decay rate is 13%; the amount of caffeine in the body decreases by 13% per hour.

3. To find the time at which only 20 mg of caffeine remains in the body, substitute  $y = 20$  and solve for the corresponding value of  $t$ .

$$20 = 120(0.87)^t$$

Divide both sides by 120 to isolate the exponential expression.

$$\frac{20}{120} = (0.87)^t$$

$$0.1667 = (0.87)^t$$

Rewrite the expression in logarithmic form and then use the change of base formula to evaluate.

$$t = \log_{0.87}(0.1667)$$

$$t = \frac{\ln(0.1667)}{\ln(0.87)}$$

$$t \approx 12.9 \text{ hours}$$

After 12.9 hours, 20 mg of caffeine remains in the body.

### 5.5.2 Expressing Exponential Functions in Different Forms

Now that we've developed our equation solving skills, we revisit the question of expressing exponential functions equivalently in the forms  $y = ab^t$  and  $y = ae^{kt}$ .

We've already determined that if given the form  $y = ae^{kt}$ , it is straightforward to find  $b$ .

**Example 5.5.7.** *For the following examples, assume  $t$  is measured in years.*

1. Express  $y = 3500e^{0.25t}$  in form  $y = ab^t$  and find the annual percentage growth rate.
2. Express  $y = 28000e^{-0.32t}$  in form  $y = ab^t$  and find the annual percentage decay rate.

**Solution 5.5.7.**

1. To express the function in the form  $y = ab^t$ : We start with the equation

$$y = 3500e^{0.25t}.$$

This can also be written as

$$y = a(e^{kt}) \quad \text{where} \quad e^{kt} = b.$$

Thus,

$$e^k = b.$$

For this example,

$$b = e^{0.25} \approx 1.284.$$

We rewrite the growth function as

$$y = 3500(1.284)^t.$$

To find  $r$ , recall that

$$b = 1 + r.$$

So,

$$1.284 = 1 + r$$

which gives us

$$r = 0.284.$$

The continuous growth rate is  $k = 0.25$  and the annual percentage growth rate is 28.4%.

2. To express  $y = 28000e^{-0.32t}$  in the form  $y = ab^t$ : We start with the equation

$$y = ab^t.$$

This can also be written as

$$y = a(e^{kt}) \quad \text{where} \quad e^{kt} = b.$$

Thus,

$$e^k = b.$$

For this example,

$$b = e^{-0.32} \approx 0.7261.$$

We rewrite the decay function as

$$y = 28000(0.7261)^t.$$

To find  $r$ , recall that

$$b = 1 + r.$$

So,

$$0.7261 = 1 + r$$

which gives us

$$r = -0.2739.$$

The continuous decay rate is  $k = -0.32$  and the annual percentage decay rate is 27.39%.

In the sentence, we omit the negative sign when stating the annual percentage decay rate because we have used the word "decay" to indicate that  $r$  is negative.

**Example 5.5.8.** Express the following:

1.  $y = 4200(1.078)^t$  in the form  $y = ae^{kt}$
2.  $y = 150(0.73)^t$  in the form  $y = ae^{kt}$

**Solution 5.5.8.** 1. To express  $y = 4200(1.078)^t$  in the form  $y = ae^{kt}$ :

Start with the equation

$$y = ae^{kt}$$

which can also be written as

$$y = ab^t.$$

We need to find the constant  $k$  such that

$$e^k = b.$$

Given  $b = 1.078$ , we have

$$e^k = 1.078.$$

Therefore,

$$k = \ln(1.078) \approx 0.0751.$$

We rewrite the growth function as

$$y = 4200e^{0.0751t}.$$

2. To express  $y = 150(0.73)^t$  in the form  $y = ae^{kt}$ : Start with the equation

$$y = ae^{kt}$$

which can also be written as

$$y = ab^t.$$

Here, we need to find the constant  $k$  such that

$$e^k = b.$$

Given  $b = 0.73$ , we have

$$e^k = 0.73.$$

Therefore,

$$k = \ln(0.73) \approx -0.3147.$$

We rewrite the decay function as

$$y = 150e^{-0.3147t}.$$

**Example 5.5.9.** Suppose that Vinh invests \$10000 in an investment earning 5% per year. He wants to know how long it will take his investment to accumulate to \$12000, and how long it would take to accumulate to \$15000.

**Solution 5.5.9.** We start by writing the exponential growth function that models the value of this investment as a function of the time since the \$10000 is initially invested

$$y = 10000(1.05)^t$$

We divide both sides by 10000 to isolate the exponential expression on one side.

$$\frac{y}{10000} = 1.05^t$$

Next we take the natural logarithm of both sides of the equation.

$$\ln(1.05^t) = \ln\left(\frac{y}{10000}\right)$$

Use the exponent property of logarithms to get the power out of the logarithm

$$t \ln(1.05) = \ln\left(\frac{y}{10000}\right)$$

Dividing both sides by  $\ln(1.05)$

$$t = \frac{\ln\left(\frac{y}{10000}\right)}{\ln(1.05)}$$

To find the number of years until the value of this investment is \$12000, we substitute  $y = 12000$  into  $y$  and evaluate to find  $t$ :

$$t = \frac{\ln\left(\frac{12000}{10000}\right)}{\ln(1.05)} = \frac{\ln(1.2)}{\ln(1.05)} \approx 3.74 \text{ years}$$

To find the number of years until the value of this investment is \$15000, we substitute  $y = 15000$  into  $y$  and evaluate to find  $t$ :

$$t = \frac{\ln\left(\frac{15000}{10000}\right)}{\ln(1.05)} = \frac{\ln(1.5)}{\ln(1.05)} \approx 8.31 \text{ years}$$

# Chapter 6

## Mathematics of Finance

In this chapter, you will learn to:

1. Solve financial problems that involve simple interest.
2. Solve problems involving compound interest.
3. Find the future value of an annuity, and the amount of payments to a sinking fund.
4. Find the future value of an annuity, and an installment payment on a loan.

### 6.1 Simple Interest and Discount

In this section, you will learn to:

1. Find simple interest.
2. Find present value.
3. Find discounts and proceeds.

#### 6.1.1 Simple Interest

It costs to borrow money. The rent one pays for the use of money is called the interest. The amount of money that is being borrowed or loaned is called

the principal or present value. Simple interest is paid only on the original amount borrowed. When the money is loaned out, the person who borrows the money generally pays a fixed rate of interest on the principal for the time period he keeps the money. Although the interest rate is often specified for a year, it may be specified for a week, a month, or a quarter, etc. The credit card companies often list their charges as monthly rates, sometimes it is as high as 1.5% a month.

### Summary 6.1.1: Simple Interest

If an amount  $P$  is borrowed for a time  $t$  at an interest rate of  $r$  per time period, then the simple interest is given by

$$I = P \cdot r \cdot t$$

The total amount  $A$ , also called the accumulated value or the future value, is given by

$$A = P + I = P + Prt$$

or

$$A = P(1 + rt)$$

where interest rate  $r$  is expressed in decimals.

**Example 6.1.1.** *Ursula borrows \$600 for 5 months at a simple interest rate of 15% per year. Find the interest, and the total amount she is obligated to pay?*

**Solution 6.1.1.** *The interest is computed by multiplying the principal with the interest rate and the time.*

$$I = Prt$$

$$I = \$600(0.15)\frac{5}{12} = \$37.50$$

*The total amount is*

$$A = P + I = \$600 + \$37.50 = \$637.50$$

*Incidentally, the total amount can be computed directly as*

$$A = P(1 + rt) = \$600[1 + (0.15)(5/12)] = \$600(1 + 0.0625) = \$637.50$$



**Example 6.1.2.** Jose deposited \$2500 in an account that pays 6% simple interest. How much money will he have at the end of 3 years?

**Solution 6.1.2.** The total amount or the future value is given by  $A = P(1 + rt)$ .

$$A = \$2500[1 + (0.06)(3)]$$

$$A = \$2950$$

**Example 6.1.3.** Darnel owes a total of \$3060 which includes 12% interest for the three years he borrowed the money. How much did he originally borrow?

**Solution 6.1.3.** This time we are asked to compute the principal  $P$ .

$$\$3060 = P[1 + (.12)(3)]$$

$$\$3060 = P(1.36)$$

$$P = \frac{\$3060}{1.36}$$

$$P = \$2250$$

Darnel originally borrowed \$2250.

**Example 6.1.4.** A Visa credit card company charges a 1.5% finance charge each month on the unpaid balance. If Martha owed \$2350 and has not paid her bill for three months, how much does she owe now?

**Solution 6.1.4.** Before we attempt the problem, the reader should note that in this problem the rate of finance charge is given per month and not per year.

The total amount Martha owes is the previous unpaid balance plus the finance charge.

$$A = \$2350 + \$2350(0.015)(3) = \$2350 + \$105.75 = \$2455.75$$

Alternatively, again, we can compute the amount directly by using formula  $A = P(1 + rt)$

$$A = \$2350[1 + (.015)(3)] = \$2350(1.045) = \$2455.75$$

### 6.1.2 Discounts and Proceeds

Banks often deduct the simple interest from the loan amount at the time that the loan is made. When this happens, we say the loan has been discounted. The interest that is deducted is called the discount, and the actual amount that is given to the borrower is called the proceeds. The amount the borrower is obligated to repay is called the maturity value.

#### Summary 6.1.2: Discounts and Proceeds

If an amount  $M$  is borrowed for a time  $t$  at a discount rate of  $r$  per year, then the discount  $D$  is

$$D = M \cdot r \cdot t$$

The proceeds  $P$ , the actual amount the borrower gets, is given by

$$P = M - D$$

$$P = M - Mrt$$

or

$$P = M(1 - rt)$$

where interest rate  $r$  is expressed in decimals.

**Example 6.1.5.** *Francisco borrows \$1200 for 10 months at a simple interest rate of 15% per year. Determine the discount and the proceeds.*

**Solution 6.1.5.** *The discount  $D$  is the interest on the loan that the bank deducts from the loan amount.*

$$D = Mrt$$

$$D = \$1200 (0.15) \left( \frac{10}{12} \right) = \$150$$

*Therefore, the bank deducts \$150 from the maturity value of \$1200, and gives Francisco \$1050. Francisco is obligated to repay the bank \$1200. In this case, the discount  $D = \$150$ , and the proceeds  $P = \$1200 - \$150 = \$1050$ .*

**Example 6.1.6.** *If Francisco wants to receive \$1200 for 10 months at a simple interest rate of 15% per year, what amount of loan should he apply for?*

**Solution 6.1.6.** *In this problem, we are given the proceeds  $P$  and are being asked to find the maturity value  $M$ .*

$$P = \$1200, \quad r = 0.15, \quad t = \frac{10}{12}$$

*. We need to find  $M$ . We know  $P = M - D$  but also  $D = Mrt$  therefore  $P = M - Mrt = M(1 - rt)$*

$$\$1200 = M \left[ 1 - (0.15) \left( \frac{10}{12} \right) \right]$$

*We need to solve for  $M$ .*

$$\$1200 = M(1 - 0.125)$$

$$\$1200 = M(0.875)$$

$$\frac{\$1200}{0.875} = M$$

$$\$1371.43 = M$$

*Therefore, Francisco should ask for a loan for \$1371.43. The bank will discount \$171.43 and Francisco will receive \$1200.*

## 6.2 Compound Interest

In this section you will learn to:

1. Find the future value of a lump-sum.
2. Find the present value of a lump-sum.
3. Find the effective interest rate.

### 6.2.1 Compound Interest

In the last section, we examined problems involving simple interest. Simple interest is generally charged when the lending period is short and often less than a year. When the money is loaned or borrowed for a longer time period, if the interest is paid (or charged) not only on the principal, but also on the past interest, then we say the interest is compounded.

Suppose we deposit \$200 in an account that pays 8% interest. At the end of one year, we will have  $\$200 + \$200(0.08) = \$200(1 + .08) = \$216$ .

Now suppose we put this amount, \$216, in the same account. After another year, we will have  $\$216 + \$216(0.08) = \$216(1 + .08) = \$233.28$ .

An initial deposit of \$200 has accumulated to \$233.28 in two years. Further note that had it been simple interest, this amount would have accumulated to only \$232. The reason the amount is slightly higher is because the interest (\$16) we earned the first year, was put back into the account. And this \$16 amount itself earned interest of  $\$16(0.08) = \$1.28$ , thus resulting in the increase. So we have earned interest on the principal as well as on the past interest, and that is why we call it compound interest.

Now suppose we leave this amount, \$233.28, in the bank for another year, the final amount will be  $\$233.28 + \$233.28(0.08) = \$233.28(1 + .08) = \$251.94$ .

Now let us look at the mathematical part of this problem so that we can devise an easier way to solve these problems. After one year, we had  $\$200(1 + .08) = \$216$ . After two years,  $\$216 = \$200(1 + .08)^2$ . But  $\$216 = \$200(1 + .08)$ , therefore, the above expression becomes

$$\$200(1 + .08)(1 + .08) = \$200(1 + .08)^2 = \$233.28$$

After three years, we get

$$\$233.28(1 + .08) = \$200(1 + .08)(1 + .08)(1 + .08)$$

which can be written as

$$\$200(1 + .08)^3 = \$251.94$$

Suppose we are asked to find the total amount at the end of 5 years, we will get

$$200(1 + .08)^5 = \$293.87$$

We summarize the compound interest calculations as follows:

The original amount	\$200	= \$200
The amount after one year	$\$200(1 + .08)$	= \$216
The amount after two years	$\$200(1 + .08)^2$	= \$233.28
The amount after three years	$\$200(1 + .08)^3$	= \$251.94
The amount after five years	$\$200(1 + .08)^5$	= \$293.87
The amount after t years	$\$200(1 + .08)^t$	

### 6.2.2 Compounding Periods

Banks often compound interest more than one time a year. Consider a bank that pays 8% interest but compounds it four times a year, or quarterly. This means that every quarter the bank will pay an interest equal to one-fourth of 8%, or 2%.

Now if we deposit \$200 in the bank, after one quarter we will have  $\$200(1 + .08/4)$  or \$204. After two quarters, we will have  $\$200(1 + .08/4)^2$  or \$208.08. After one year, we will have  $\$200(1 + .08/4)^4$  or \$216.49. After three years, we will have  $\$200(1 + .08/4)^{12}$  or \$253.65, etc.

The original amount	\$200	= \$200
The amount after one quarter	$\$200 \left(1 + \frac{.08}{4}\right)$	= \$204
The amount after two quarters	$\$200 \left(1 + \frac{.08}{4}\right)^2$	= \$208.08
The amount after one year	$\$200 \left(1 + \frac{.08}{4}\right)^4$	= \$216.49
The amount after two years	$\$200 \left(1 + \frac{.08}{4}\right)^8$	= \$234.31
The amount after three years	$\$200 \left(1 + \frac{.08}{4}\right)^{12}$	= \$253.65
The amount after five years	$\$200 \left(1 + \frac{.08}{4}\right)^{20}$	= \$297.19
The amount after t years	$\$200 \left(1 + \frac{.08}{4}\right)^{4t}$	

The general formula for compound interest is given by the

#### Summary 6.2.1: Compound Interest Formula

$$A = P \left(1 + \frac{r}{n}\right)^{nt}$$

where  $P$  is the principal amount,  $r$  is the annual interest rate,  $n$  is the number of times interest is compounded per year, and  $t$  is the time in years.

**Example 6.2.1.** If \$3500 is invested at 9% compounded monthly, what will

the future value be in four years?

**Solution 6.2.1.** Clearly an interest of  $\frac{0.09}{12}$  is paid every month for four years. The interest is compounded  $4 \cdot 12 = 48$  times over the four-year period. We get

$$A = \$3500 \left(1 + \frac{.09}{12}\right)^{48} = \$3500(1.0075)^{48} = \$5009.92$$

\$3500 invested at 9% compounded monthly will accumulate to \$5009.92 in four years.

**Example 6.2.2.** How much should be invested in an account paying 9% compounded daily for it to accumulate to \$5,000 in five years?

**Solution 6.2.2.** We know the future value, but need to find the principal.

$$\$5000 = P \left(1 + \frac{.09}{365}\right)^{365(5)}$$

$$\$5000 = P(1.568225)$$

$$P = \$3188.32$$

\$3188.32 invested in an account paying 9% compounded daily will accumulate to \$5,000 in five years.

**Example 6.2.3.** If \$4,000 is invested at 4% compounded annually, how long will it take to accumulate to \$6,000?

**Solution 6.2.3.**  $n = 1$  because annual compounding means compounding only once per year. The formula simplifies to  $A = (1 + r)^t$  when  $n = 1$ .

$$\$6000 = \$4000(1.04)^t$$

Dividing by 4000 yields

$$\frac{6000}{4000} = (1.04)^t$$

$$1.5 = 1.04^t$$

We use logarithms to solve for the value of  $t$  because the variable  $t$  is in the exponent.

$$\log(1.5) = \log(1.04^t)$$

Using the Exponent Rule for logarithms we can take the power down

$$\log(1.5) = t \log(1.04)$$

then solve for  $t$  by dividing:

$$t = \frac{\ln(1.5)}{\ln(1.04)} \approx 10.33 \text{ years}$$

It takes about 10 years and 4 months for \$4000 to accumulate to \$6000 if invested at 4% interest, compounded annually.

**Example 6.2.4.** If \$5,000 is invested now for 6 years what interest rate compounded quarterly is needed to obtain an accumulated value of \$8,000.

**Solution 6.2.4.** We have  $n = 4$  for quarterly compounding.

$$\begin{aligned} \$8,000 &= \$5,000 \left(1 + \frac{r}{4}\right)^{4 \cdot 6} \\ \$8,000 &= \$5,000 \left(1 + \frac{r}{4}\right)^{24} \\ 1.6 &= \left(1 + \frac{r}{4}\right)^{24} \end{aligned}$$

We use roots to solve for  $r$  because the variable  $r$  is in the base, whereas the exponent is a known number.

$$\begin{aligned} \sqrt[24]{1.6} &= 1 + \frac{r}{4} \\ 1.6^{\frac{1}{24}} &= 1 + \frac{r}{4} \end{aligned}$$

Evaluating the left side of the equation gives

$$\begin{aligned} 1.0197765 &= 1 + \frac{r}{4} \\ 0.0197765 &= \frac{r}{4} \\ r &= 4(0.0197765) = 0.0791 \end{aligned}$$

An interest rate of 7.91% is needed in order for \$5,000 invested now to accumulate to \$8,000 at the end of 6 years, with interest compounded quarterly.

### 6.2.3 Effective Interest Rate

Banks are required to state their interest rate in terms of an “effective yield” or “effective interest rate”, for comparison purposes. The effective rate is also called the Annual Percentage Yield (APY) or Annual Percentage Rate (APR).

The effective rate is the interest rate compounded annually would be equivalent to the stated rate and compounding periods. The next example shows how to calculate the effective rate. To examine several investments to see which has the best rate, we find and compare the effective rate for each investment.

Example 6.2.5 illustrates how to calculate the effective rate.

**Example 6.2.5.** *If Bank A pays 7.2% interest compounded monthly, what is the effective interest rate? If Bank B pays 7.25% interest compounded semiannually, what is the effective interest rate? Which bank pays more interest?*

**Solution 6.2.5.** *Bank A: Suppose we deposit \$1 in this bank and leave it for a year, we will get*

$$r_{effective} = \left(1 + \frac{0.072}{12}\right)^{12} = 1.0744$$

$$r_{effective} = 1.0744 - 1 = 0.0744$$

*We earned interest of \$1.0744 - \$1.00 = \$0.0744 on an investment of \$1.*

*The effective interest rate is 7.44%, often referred to as the APY or APR.*

*Bank B: The effective rate is calculated as*

$$r_{effective} = \left(1 + \frac{0.0725}{2}\right)^2 - 1 = .0738$$

*The effective interest rate is 7.38%.*

*Bank A pays slightly higher interest, with an effective rate of 7.44%, compared to Bank B with effective rate 7.38%.*



### 6.2.4 Continuous Compounding

Interest can be compounded yearly, semiannually, quarterly, monthly, and daily. Using the same calculation methods, we could compound every hour, every minute, and even every second. As the compounding period gets shorter and shorter, we move toward the concept of continuous compounding.

But what do we mean when we say the interest is compounded continuously, and how do we compute such amounts? When interest is compounded "infinitely many times", we say that the interest is compounded continuously. Our next objective is to derive a formula to model continuous compounding.

Suppose we put \$1 in an account that pays 100% interest. If the interest is compounded once a year, the total amount after one year will be  $\$1(1 + 1) = \$2$ . If the interest is compounded semiannually, in one year we will have  $\$1(1 + \frac{1}{2})^2 = \$2.25$ . If the interest is compounded quarterly, in one year we will have  $\$1(1 + \frac{1}{4})^4 = \$2.44$ . If the interest is compounded monthly, in one year we will have  $\$1(1 + \frac{1}{12})^{12} = \$2.61$ . If the interest is compounded daily, in one year we will have  $\$1(1 + \frac{1}{365})^{365} = \$2.71$ .

We show the results as follows:

Frequency of compounding	Formula	Total amount
Annually	$\$1(1 + 1)$	\$2
Semiannually	$\$1(1 + \frac{1}{2})^2$	\$2.25
Quarterly	$\$1(1 + \frac{1}{4})^4$	\$2.44140625
Monthly	$\$1(1 + \frac{1}{12})^{12}$	\$2.61303529
Daily	$\$1(1 + \frac{1}{365})^{365}$	\$2.71456748
Hourly	$\$1(1 + \frac{1}{8760})^{8760}$	\$2.71812699
Every minute	$\$1(1 + \frac{1}{525600})^{525600}$	\$2.71827922
Every Second	$\$1(1 + \frac{1}{31536000})^{31536000}$	\$2.71828247
Continuously	$\$1(2.718281828\dots)$	\$2.718281828...

We have noticed that the \$1 we invested does not grow without bound. It starts to stabilize to an irrational number 2.718281828... given the name "e" after the great mathematician Euler. In mathematics, we say that as  $n$  becomes infinitely large, the expression equals  $e$ . Therefore, it is natural that the number  $e$  plays a part in continuous compounding.

It can be shown that as  $n$  becomes infinitely large, the expression

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$$

Therefore, it follows that if we invest \$P at an interest rate  $r$  per year, compounded continuously, after  $t$  years the final amount will be given by

**Summary 6.2.2: Continuous Compounding**

$$A = P \cdot e^{rt}$$

with  $P$  the principal,  $r$  the interest rate,  $t$  the time in years, and  $A$  the value after compounding for those  $t$  years

**Example 6.2.6.** \$3500 is invested at 9% compounded continuously. Find the future value in 4 years.

**Solution 6.2.6.** Using the formula for the continuous compounding, we get  $A = Pe^{rt}$ .

$$A = \$3500e^{0.09(4)}$$

$$A = \$3500e^{0.36}$$

$$A = \$5016.65$$

**Example 6.2.7.** If an amount is invested at 7% compounded continuously, what is the effective interest rate?

**Solution 6.2.7.** If we deposit \$1 in the bank at 7% compounded continuously for one year, and subtract that \$1 from the final amount, we get the effective interest rate in decimals.

$$r_{\text{effective}} = 1e^{0.07} - 1$$

$$r_{\text{effective}} = 1.0725 - 1$$

$$r_{\text{effective}} = 0.0725 \quad \text{or} \quad 7.25\%$$

**Example 6.2.8.** If an amount is invested at 7% compounded continuously, how long will it take to double?

**Solution 6.2.8.** We use the model:  $Pe^{0.07t} = A$ .

*We don't know the initial value of the principal but we do know that the accumulated value is double (twice) the principal.*

$$\begin{aligned}
 Pe^{0.07t} &= 2P \\
 \frac{Pe^{0.07t}}{P} &= \frac{2P}{P} \\
 e^{0.07t} &= 2 \\
 0.07t &= \ln(2) \\
 t &= \frac{\ln(2)}{0.07} \\
 t &= 9.9 \text{ years}
 \end{aligned}$$

*It takes 9.9 years for money to double if invested at 7% continuous interest.*

## 6.3 Annuities and Sinking Funds

In this section, you will learn to:

1. Find the future value of an annuity.
2. Find the amount of payments to a sinking fund.

### 6.3.1 Ordinary Annuity

In the first two sections of this chapter, we examined problems where an amount of money was deposited lump sum in an account and was left there for the entire time period. Now we will do problems where timely payments are made in an account. When a sequence of payments of some fixed amount are made in an account at equal intervals of time, we call that an annuity. And this is the subject of this section.

To develop a formula to find the value of an annuity, we will need to recall the formula for the sum of a geometric series.

A geometric series is of the form:  $a + ax + ax^2 + ax^3 + \dots + ax^n$ .

In a geometric series, each subsequent term is obtained by multiplying the preceding term by a number, called the common ratio. A geometric series is

completely determined by knowing its first term, the common ratio, and the number of terms.

The following are some examples of geometric series:

- $3 + 6 + 12 + 24 + 48$  has first term  $a = 3$  and common ratio  $x = 2$ .
- $2 + 6 + 18 + 54 + 162$  has first term  $a = 2$  and common ratio  $x = 3$ .
- $37 + 3.7 + .37 + .037 + .0037$  has first term  $a = 35$  and common ratio  $x = 0.1$ .

In your algebra class, you developed a formula for finding the sum of a geometric series. You probably used  $r$  as the symbol for the ratio, but we are using  $x$  because  $r$  is the symbol we have been using for the interest rate.

The formula for the sum of a geometric series with first term  $a$  and common ratio  $x$  is:

$$S = a \left( \frac{1 - x^n}{1 - x} \right)$$

We will use this formula to find the value of an annuity.

Consider the following example.

**Example 6.3.1.** *If at the end of each month a deposit of \$500 is made in an account that pays 8% compounded monthly, what will the final amount be after five years?*

**Solution 6.3.1.** *There are 60 deposits made in this account. The first payment stays in the account for 59 months, the second payment for 58 months, the third for 57 months, and so on.*

*The first payment of \$500 will accumulate to an amount of*

$$\$500 \left( 1 + \frac{0.08}{12} \right)^{59}.$$

*The second payment of \$500 will accumulate to an amount of*

$$\$500 \left( 1 + \frac{0.08}{12} \right)^{58}.$$

The third payment will accumulate to

$$\$500 \left(1 + \frac{0.08}{12}\right)^{57}.$$

And so on...

Finally, the next to last (59th) payment will accumulate to

$$\$500 \left(1 + \frac{0.08}{12}\right)^1$$

. The last payment is taken out the same time it is made, and will not earn any interest.

To find the total amount in five years, we need to add the accumulated value of these sixty payments.

In other words, we need to find the sum of the following series:

$$\$500 \left(1 + \frac{0.08}{12}\right)^{59} + \$500 \left(1 + \frac{0.08}{12}\right)^{58} + \$500 \left(1 + \frac{0.08}{12}\right)^{57} + \dots + \$500$$

Written backwards, we have

$$\$500 + \$500 \left(1 + \frac{0.08}{12}\right) + \$500 \left(1 + \frac{0.08}{12}\right)^2 + \dots + \$500 \left(1 + \frac{0.08}{12}\right)^{59}$$

This is a geometric series with  $a = \$500$ ,  $r = \left(1 + \frac{0.08}{12}\right)$ , and  $n = 59$ . The sum is

$$\begin{aligned} & \$500 \left( \frac{\left(1 + \frac{0.08}{12}\right)^{60} - 1}{\frac{0.08}{12}} \right) \\ &= \$500(73.47686) \\ &= \$36738.43 \end{aligned}$$

When the payments are made at the end of each period rather than at the beginning, we call it an ordinary annuity.

**Summary 6.3.1: Future Value of an Ordinary Annuity**

If a payment of  $m$  dollars is made in an account  $n$  times a year at an interest rate  $r$ , then the final amount  $A$  after  $t$  years is given by:

$$A = \frac{m \left[ \left( 1 + \frac{r}{n} \right)^{nt} - 1 \right]}{\frac{r}{n}}$$

The future value is also called the accumulated value. Note that the formula assumes that the payment period is the same as the compounding period. If these are not the same, then this formula does not apply

**Example 6.3.2.** *Tanya deposits \$300 at the end of each quarter in her savings account. If the account earns 5.75% compounded quarterly, how much money will she have in 4 years?*

**Solution 6.3.2.** *The future value of this annuity can be found using the above formula.*

$$\begin{aligned} A &= \$300 \left( \frac{\left( 1 + \frac{0.0575}{4} \right)^{4 \cdot 4} - 1}{\frac{0.0575}{4}} \right) \\ A &= \$300(17.8463) \\ A &= \$5353.89 \end{aligned}$$

*If Tanya deposits \$300 into a savings account earning 5.75% compounded quarterly for 4 years, then at the end of 4 years she will have \$5,353.89.*

**Example 6.3.3.** *Robert needs \$5,000 in three years. How much should he deposit each month in an account that pays 8% compounded monthly in order to achieve his goal?*

**Solution 6.3.3.** *If Robert saves  $m$  dollars per month, after three years he will have*

$$m \left( \frac{\left( 1 + \frac{0.08}{12} \right)^{36} - 1}{\frac{0.08}{12}} \right)$$

*But we'd like this amount to be \$5,000. Therefore,*

$$m \left( \frac{\left( 1 + \frac{0.08}{12} \right)^{36} - 1}{\frac{0.08}{12}} \right) = \$5000$$

$$m(40.5356) = \$5000$$

$$m = \frac{\$5000}{40.5356} = \$123.35$$

Robert needs to deposit \$123.35 at the end of each month for 3 years into an account paying 8% compounded monthly in order to have \$5,000 at the end of 5 years.

### 6.3.2 Sinking Funds

When a business deposits money at regular intervals into an account in order to save for a future purchase of equipment, the savings fund is referred to as a "sinking fund". Calculating the sinking fund deposit uses the same method as the previous problem.

**Example 6.3.4.** *A business needs \$450,000 in five years. How much should be deposited each quarter in a sinking fund that earns 9% compounded quarterly to have this amount in five years?*

**Solution 6.3.4.** *Again, suppose that  $m$  dollars are deposited each quarter in the sinking fund. After five years, the future value of the fund should be \$450,000. This suggests the following relationship:*

$$m \left[ \frac{(1 + \frac{0.09}{4})^{20} - 1}{\frac{0.09}{4}} \right] = \$450,000$$

$$m(24.9115) = 450,000$$

$$m = \frac{450000}{24.9115} = \$18,063.93$$

The business needs to deposit \$18,063.93 at the end of each quarter for 5 years into a sinking fund earning interest of 9% compounded quarterly in order to have \$450,000 at the end of 5 years.

### 6.3.3 Annuities Due

If the payment is made at the beginning of each period, rather than at the end, we call it an annuity due. The formula for the annuity due can be derived in a similar manner. Reconsider Example 6.3.1, with the change that the deposits are made at the beginning of each month.

**Example 6.3.5.** *If at the beginning of each month a deposit of \$500 is made in an account that pays 8% compounded monthly, what will the final amount be after five years?*

**Solution 6.3.5.** *There are 60 deposits made in this account. The first payment stays in the account for 60 months, the second payment for 59 months, the third for 58 months, and so on.*

*The first payment of \$500 will accumulate to an amount of*

$$\$500 \left(1 + \frac{0.08}{12}\right)^{59}.$$

*The second payment of \$500 will accumulate to an amount of*

$$\$500 \left(1 + \frac{0.08}{12}\right)^{58}.$$

*The third payment will accumulate to*

$$\$500 \left(1 + \frac{0.08}{12}\right)^{57}.$$

*And so on...*

*Finally, the last (60th) payment will accumulate a month's interest to*

$$\$500 \left(1 + \frac{0.08}{12}\right)^1$$

.

*To find the total amount in five years, we need to find the sum of the series:  
In other words, we need to find the sum of the following series:*

$$\$500 \left(1 + \frac{0.08}{12}\right)^{60} + \$500 \left(1 + \frac{0.08}{12}\right)^{59} + \$500 \left(1 + \frac{0.08}{12}\right)^{58} + \dots + \$500 \left(1 + \frac{0.08}{12}\right)^1$$

*Written backwards, we have*

$$\$500 \left(1 + \frac{0.08}{12}\right) + \$500 \left(1 + \frac{0.08}{12}\right)^2 + \dots + \$500 \left(1 + \frac{0.08}{12}\right)^{60}$$



#### 6.4. PRESENT VALUE OF AN ANNUITY AND INSTALLMENT PAYMENT 169

*This isn't a geometric series, but if we add \$500 to the front of the series and subtract it from the back we haven't changed the value, but we'll have a geometric series.*

$$\$500 + \$500 \left(1 + \frac{0.08}{12}\right) + \$500 \left(1 + \frac{0.08}{12}\right)^2 + \dots + \$500 \left(1 + \frac{0.08}{12}\right)^{60} - \$500$$

*Except for the last term, we have a geometric series with  $a = \$500$ ,  $r = (1 + .08/12)$ , and  $n = 60$ . Therefore the sum is*

$$A = \frac{\$500[(1 + .08/12)^{61} - 1]}{.08/12} - \$500$$

$$A = \$500(74.9667) - \$500$$

$$A = \$37483.35 - \$500$$

$$A = \$36983.35$$

So, in the case of an annuity due, to find the future value, we increase the number of periods  $n$  by 1, and subtract one payment.

##### Summary 6.3.2: Future Value of an Annuity Due

If a payment of  $m$  dollars is made in an account  $n$  times a year (at the beginning of each month) at an interest rate  $r$ , then the final amount  $A$  after  $t$  years is given by:

$$A = \frac{m \left[ \left(1 + \frac{r}{n}\right)^{nt+1} - 1 \right]}{\frac{r}{n}} - m$$

Most of the problems we are going to do in this chapter involve ordinary annuities, therefore, we will down play the significance of the last formula for the annuity due. We mentioned the formula for the annuity due only for completeness.

## 6.4 Present Value of an Annuity and Installment Payment

In this section, you will learn to:

1. Find the present value of an annuity.
2. Find the amount of installment payment on a loan.

In Section 6.2, we learned to find the future value of a lump sum, and in Section 6.3, we learned to find the future value of an annuity. With these two concepts in hand, we will now learn to amortize a loan, and to find the present value of an annuity.

The present value of an annuity is the amount of money we would need now in order to be able to make the payments in the annuity in the future. In other word, the present value is the value now of a future stream of payments.

We start by breaking this down step by step to understand the concept of the present value of an annuity. After that, the examples provide a more efficient way to do the calculations by working with concepts and calculations we have already explored in Sections 6.2 and 6.3.

Suppose Carlos owns a small business and employs an assistant manager to help him run the business. Assume it is January 1 now. Carlos plans to pay his assistant manager a \$1000 bonus at the end of this year and another \$1000 bonus at the end of the following year. Carlos' business had good profits this year so he wants to put the money for his assistant's future bonuses into a savings account now. The money he puts in now will earn interest at the rate of 4% per year compounded annually while in the savings account.

How much money should Carlos put into the savings account now so that he will be able to withdraw \$1000 one year from now and another \$1000 two years from now?

At first, this sounds like a sinking fund. But it is different. In a sinking fund, we put money into the fund with periodic payments to save to accumulate to a specified lump sum that is the future value at the end of a specified time period.

In this case we want to put a lump sum into the savings account now, so that lump sum is our principal,  $P$ . Then we want to withdraw that amount as a series of period payments; in this case the withdrawals are an annuity with \$1000 payments at the end of each two years.

We need to determine the amount we need in the account now, the present value, to be able to make withdraw the periodic payments later.

#### 6.4. PRESENT VALUE OF AN ANNUITY AND INSTALLMENT PAYMENT 171

We use the compound interest formula from Section 6.2 with  $r = 0.04$  and  $n = 1$  for annual compounding to determine the present value of each payment of \$1000.

Consider the first payment of \$1000 at the end of year 1. Let  $P_1$  be its present value

$$\$1000 = P(1.04)^1 \quad \text{so} \quad P_1 = \$961.54$$

Now consider the second payment of \$1000 at the end of year 2. Let  $P_2$  be its present value

$$\$1000 = P(1.04)^2 \quad \text{so} \quad P_2 = \$924.56$$

To make the \$1000 payments at the specified times in the future, the amount that Carlos needs to deposit now is the present value.

The calculation above was useful to illustrate the meaning of the present value of an annuity.

But it is not an efficient way to calculate the present value. If we were to have a large number of annuity payments, the step-by-step calculation would be long and tedious.

Example 6.4.1 investigates and develops an efficient way to calculate the present value of an annuity, by relating the future (accumulated) value of an annuity and its present value.

**Example 6.4.1.** *Suppose you have won a lottery that pays \$1,000 per month for the next 20 years. But, you prefer to have the entire amount now. If the interest rate is 8%, how much will you accept?*

**Solution 6.4.1.** *This classic present value problem needs our complete attention because the rationalization we use to solve this problem will be used again in the problems to follow.*

*Consider, for argument purposes, that two people Mr. Cash, and Mr. Credit have won the same lottery of \$1,000 per month for the next 20 years. Mr. Credit is happy with his \$1,000 monthly payment, but Mr. Cash wants to have the entire amount now.*

*Our job is to determine how much Mr. Cash should get. We reason as follows:*

*If Mr. Cash accepts \$P dollars, then the \$P dollars deposited at 8% for 20 years should yield the same amount as the \$1,000 monthly payments for 20 years. In other words, we are comparing the future values for both Mr. Cash and Mr. Credit, and we would like the future values to equal.*

*Since Mr. Cash is receiving a lump sum of x dollars, its future value is given by the lump sum formula we studied in Section 6.2, and it is*

$$A = P(1 + 0.08/12)^{240}$$

*Since Mr. Credit is receiving a sequence of payments, or an annuity, of \$1,000 per month, its future value is given by the annuity formula we learned in Section 6.3. This value is*

$$A = \frac{\$1,000[(1 + 0.08/12)^{240} - 1]}{0.08/12}$$

*The only way Mr. Cash will agree to the payment he receives is if these two future values are equal. So we set them equal and solve for the unknown.*

$$P(1 + 0.08/12)^{240} = \frac{\$1,000[(1 + 0.08/12)^{240} - 1]}{0.08/12}$$

$$P(4.9268) = \$1,000(589.02041)$$

$$P = \frac{\$589,020.41}{4.9268}$$

$$P = \$119,554.36$$

*The present value of an ordinary annuity of \$1,000 each month for 20 years at 8% is \$119,554.36.*

*The reader should also note that if Mr. Cash takes his lump sum of  $P = \$119,554.36$  and invests it at 8% compounded monthly, he will have an accumulated value of  $A = \$589,020.41$  in 20 years.*

### 6.4.1 Installment Payment on a Loan

If a person or business needs to buy or pay for something now (a car, a home, college tuition, equipment for a business) but does not have the money, they can borrow the money as a loan. They receive the loan amount called the principal (or present value) now and are obligated to pay back the principal in the future over a stated amount of time (term of the loan), as regular periodic payments with interest.

Example 6.4.2 examines how to calculate the loan payment, using reasoning similar to Example 6.4.1.

**Example 6.4.2.** *Find the monthly payment for a car costing \$15,000 if the loan is amortized over five years at an interest rate of 9%.*

**Solution 6.4.2.** *Consider the scenario where Mr. Cash pays cash and Mr. Credit wishes to make monthly payments. We are to determine the monthly payment amount for Mr. Credit.*

*Since Mr. Cash is paying a lump sum of \$15,000, his future value is given by the lump sum formula:*

$$A = P(1 + r/n)^{nt} = \$15,000(1 + .09/12)^{60} = \$15,000(1.5657)$$

*For Mr. Credit to make a sequence of payments, or an annuity, of  $m$  dollars per month, and its future value is given by the annuity formula:*

$$A = m \left[ \frac{(1 + r/n)^{nt} - 1}{r/n} \right] = m \left[ \frac{(1 + .09/12)^{12(5)} - 1}{.09/12} \right] = m(75.4241)$$

*Setting the two future values equal and solving for  $m$ , we find the monthly payment  $m$ :*

$$\$15,000(1.5657) = m(75.4241)$$

$$m = \frac{\$15,000 \cdot 1.5657}{75.4241} = \$311.38$$

*Therefore, the monthly payment needed to repay the loan is \$311.38 for five years.*

**Summary 6.4.1: Present Value of an Annuity**

If a payment of  $m$  dollars is made in an account  $n$  times a year at an interest  $r$ , then the present value  $P$  of the annuity after  $t$  years is

$$P(1 + r/n)^{nt} = \frac{m[(1 + r/n)^{nt} - 1]}{r/n}$$

**Summary 6.4.2: Present Value of an Annuity**

When used for a loan, the amount  $P$  is the loan amount, and  $m$  is the periodic payment needed to repay the loan over a term of  $t$  years with  $n$  payments per year at an interest  $r$ .

$$P(1 + r/n)^{nt} = \frac{m[(1 + r/n)^{nt} - 1]}{r/n}$$

Note that the formula assumes that the payment period is the same as the compounding period. If these are not the same, then this formula does not apply.

## 6.4.2 Alternate Formula to find Present Value of an Annuity

Finally, we note that many finite mathematics and finance books develop the formula for the present value of an annuity differently.

Instead of using the formula:

$$P \left(1 + \frac{r}{n}\right)^{nt} = \frac{m[(1 + \frac{r}{n})^{nt} - 1]}{\frac{r}{n}}$$

and solving for the present value  $P$  after substituting the numerical values for the other items in the formula, many textbooks first solve the formula for  $P$  in order to develop a new formula for the present value. Then the numerical information can be substituted into the present value formula and evaluated, without needing to solve algebraically for  $P$ .

Starting with Summary 6.4.1:

$$P \left(1 + \frac{r}{n}\right)^{nt} = \frac{m[(1 + \frac{r}{n})^{nt} - 1]}{\frac{r}{n}}$$

Divide both sides by  $(1 + \frac{r}{n})^{nt}$  to isolate  $P$ ,

$$P = m \frac{[(1 + \frac{r}{n})^{nt} - 1]}{\frac{r}{n}} \cdot \frac{1}{(1 + \frac{r}{n})^{nt}}$$

simplify to arrive at

**Summary 6.4.3: Present Value of an Annuity (Alternate Formula)**

$$P = m \frac{[1 - (1 + \frac{r}{n})^{-nt}]}{\frac{r}{n}}$$

The authors of this book believe that it is easier to use the formula in Sumamry 6.4.1 and solve for  $P$  or  $m$  as needed. In this approach there are fewer formulas to understand, and many students find it easier to learn. In the problems the rest of this chapter, when a problem requires the calculation of the present value of an annuity, formula in Sumamry 6.4.1 will be used.

However, some people prefer formula in Summary 6.4.3, and it is mathematically correct to use that method. Note that if you choose to use formula in Summary 6.4.3, you need to be careful with the negative exponents in the formula. And if you needed to find the periodic payment, you would still need to do the algebra to solve for the value of  $m$ .

It would be a good idea to check with your instructor to see if he or she has a preference. In fact, you can usually tell your instructor's preference by noting how he or she explains and demonstrates these types of problems in class.

## 6.5 Miscellaneous Application Problems

In this section, you will learn to apply concepts of compound interest for savings and annuities to:

1. Find the outstanding balance, partway through the term of a loan, of the future payments still remaining on the loan.
2. Perform financial calculations in situations involving several stages of savings and/or annuities.
3. Find the fair market value of a bond.
4. Construct an amortization schedule for a loan.

We have already developed the tools to solve most finance problems. Now we use these tools to solve some application problems.

### 6.5.1 Outstanding Balance on a Loan

One of the most common problems deals with finding the balance owed at a given time during the life of a loan. Suppose a person buys a house and amortizes the loan over 30 years, but decides to sell the house a few years later. At the time of the sale, he is obligated to pay off his lender; therefore, he needs to know the balance he owes. Since most long-term loans are paid off prematurely, we are often confronted with this problem.

To find the outstanding balance of a loan at a specified time, we need to find the present value  $P$  of all future payments that have not yet been paid. In this case,  $t$  does not represent the entire term of the loan. Instead:

- $t$  represents the time that still remains on the loan
- $nt$  represents the total number of future payments.

**Example 6.5.1.** *Mr. Jackson bought his house in 2004, and financed the loan for 30 years at an interest rate of 7.8%. His monthly payment was \$1260. In 2024, Mr. Jackson decides to pay off the loan. Find the balance of the loan he still owes.*

**Solution 6.5.1.** *The reader should note that the original amount of the loan is not mentioned in the problem. That is because we don't need to know that to find the balance.*

*The original loan was for 30 years. 20 years have past so there are  $10 = 30 - 20$  years still remaining.  $12(10) = 120$  payments still remain to be paid on this loan.*



As for the bank or lender is concerned, Mr. Jackson is obligated to pay \$1260 each month for 10 more years; he still owes a total of 120 payments. But since Mr. Jackson wants to pay it all off now, we need to find the present value  $P$  at the time of repayment of the remaining 10 years of payments of \$1260 each month.

Using the formula we get for the present value of an annuity, we get

$$P \left( 1 + \frac{0.078}{12} \right)^{120} = \$1260 \left[ \frac{\left( 1 + \frac{0.078}{12} \right)^{120} - 1}{\frac{0.078}{12}} \right]$$

$$P(2.17597) = \$227957.85$$

$$P = \$104761.48$$

#### Summary 6.5.1: Finding the Outstanding Balance of a Loan

If a loan has a payment of  $m$  dollars made  $n$  times a year at an interest  $r$ , then the outstanding value of the loan when there are  $t$  years still remaining on the loan is given by  $P$ :

$$P \left( 1 + \frac{r}{n} \right)^{nt} = \frac{m \left[ \left( 1 + \frac{r}{n} \right)^{nt} - 1 \right]}{\frac{r}{n}}$$

Note that  $t$  is not the original term of the loan but instead  $t$  is the amount of time still remaining in the future  $nt$  is the number of payments still remaining in the future.

Note that there are other methods to find the outstanding balance on a loan, but the method illustrated above is the easiest.

One alternate method would be to use an amortization schedule, as illustrated toward the end of this section. An amortization schedule shows the payments, interest, and outstanding balance step by step after each loan payment. An amortization schedule is tedious to calculate by hand but can be easily constructed using spreadsheet software.

Another way to find the outstanding balance, which we will not illustrate here, is to find the difference  $A - B$ , where:

- $A$  is the original loan amount (principal) accumulated to the date on which we want to find the outstanding balance (using compound interest formula).

- $B$  is the accumulated value of all payments that have been made as of the date on which we want to find the outstanding balance (using formula for the accumulated value of an annuity).

In this case, we would need to do a compound interest calculation and an annuity calculation; then, we need to find the difference between them. Three calculations are needed instead of one.

It is a mathematically acceptable way to calculate the outstanding balance. However, it is strongly recommended that students use the method explained in the box above and illustrated in Example 6.5.1, as it is much simpler.

### 6.5.2 Problems Involving Multiple Stages of Savings and/or Annuities

1. Suppose a baby, Aisha, is born and her grandparents invest \$5000 in a college fund. The money remains invested for 18 years until Aisha enters college, and then is withdrawn in equal semiannual payments over the 4 years that Aisha expects to need to finish college. The college investment fund earns 5% interest compounded semiannually. How much money can Aisha withdraw from the account every six months while she is in college?
2. Aisha graduates college and starts a job. She saves \$1000 each quarter, depositing it into a retirement savings account. Suppose that Aisha saves for 30 years and then retires. At retirement she wants to withdraw money as an annuity that pays a constant amount every month for 25 years. During the savings phase, the retirement account earns 6% interest compounded quarterly. During the annuity payout phase, the retirement account earns 4.8% interest compounded monthly. Calculate Aisha's monthly retirement annuity payout.

These problems appear complicated. But each can be broken down into two smaller problems involving compound interest on savings or involving annuities. Often the problem involves a savings period followed by an annuity period; the accumulated value from the first part of the problem may become a present value in the second part. Read each problem carefully to determine what is needed.

**Example 6.5.2.** *Suppose a baby, Aisha, is born and her grandparents invest*

\$8000 in a college fund. The money remains invested for 18 years until Aisha enters college, and then is withdrawn in equal semiannual payments over the 4 years that Aisha expects to attend college. The college investment fund earns 5% interest compounded semiannually. How much money can Aisha withdraw from the account every six months while she is in college?

**Solution 6.5.2.**

**Part 1: Accumulation of College Savings:** Find the accumulated value at the end of 18 years of a sum of \$8000 invested at 5% compounded semiannually.

$$A = \$8000 \left(1 + \frac{0.05}{2}\right)^{2(18)} = \$8000(1.025)^{36} = \$8000(2.432535) = \$19460.28$$

**Part 2: Semiannual annuity payout from savings to put toward college expenses.** Find the amount of the semiannual payout for four years using the accumulated savings from part 1 of the problem with an interest rate of 5% compounded semiannually.

$$A = \$19460.28$$

in Part 1 is the accumulated value at the end of the savings period. This becomes the present value  $P = \$19460.28$  when calculating the semiannual payments in Part 2.

$$\$19460.28 \left(1 + \frac{0.05}{2}\right)^{2(4)} = \frac{m \left[ \left(1 + \frac{0.05}{2}\right)^{2(4)} - 1 \right]}{\left(\frac{0.05}{2}\right)}$$

$$\$23710.46 = m(8.73612)$$

$$m = \$2714.07$$

Aisha will be able to withdraw \$2714.07 semiannually for her college expenses.

**Example 6.5.3.** Aisha graduates college and starts a job. She saves \$1000 each quarter, depositing it into a retirement savings account. Suppose that Aisha saves for 30 years and then retires. At retirement she wants to withdraw money as an annuity that pays a constant amount every month for 25 years. During the savings phase, the retirement account earns 6% interest

compounded quarterly. During the annuity payout phase, the retirement account earns 4.8% interest compounded monthly. Calculate Aisha's monthly retirement annuity payout.

**Solution 6.5.3.**

**Part 1: Accumulation of Retirement Savings:** Find the accumulated value at the end of 30 years of \$1000 deposited at the end of each quarter into a retirement savings account earning 6% interest compounded quarterly.

$$A = \frac{\$1000 \left[ \left( 1 + \frac{0.06}{4} \right)^{4(30)} - 1 \right]}{\left( \frac{0.06}{4} \right)}$$

$$A = \$331288.19$$

**Part 2: Monthly retirement annuity payout:** Find the amount of the monthly annuity payments for 25 years using the accumulated savings from part 1 of the problem with an interest rate of 4.8% compounded monthly.

$$A = \$331288.19$$

in Part 1 is the accumulated value at the end of the savings period. This amount will become the present value  $P = \$331288.19$  when calculating the monthly retirement annuity payments in Part 2.

$$\$331288.19 \left( 1 + \frac{0.048}{12} \right)^{12 \cdot 25} = \frac{m \left[ \left( 1 + \frac{0.048}{12} \right)^{12 \cdot 25} - 1 \right]}{\frac{0.048}{12}}$$

$$\$1097285.90 = m(578.04483)$$

$$m = \$1898.27$$

Aisha will have a monthly retirement annuity income of \$1898.27 when she retires.

### 6.5.3 Fair Market Value of a Bond

Whenever a business, and for that matter the U. S. government, needs to raise money it does it by selling bonds. A bond is a certificate of promise that states the terms of the agreement. Usually the business sells bonds for

the face amount of \$1,000 each for a stated term, a period of time ending at a specified maturity date.

The person who buys the bond, the bondholder, pays \$1,000 to buy the bond.

The bondholder is promised two things: First that he will get his \$1,000 back at the maturity date, and second that he will receive a fixed amount of interest every six months.

As the market interest rates change, the price of the bond starts to fluctuate. The bonds are bought and sold in the market at their fair market value.

The interest rate a bond pays is fixed, but if the market interest rate goes up, the value of the bond drops since the money invested in the bond could earn more if invested elsewhere. When the value of the bond drops, we say it is trading at a discount.

On the other hand, if the market interest rate drops, the value of the bond goes up since the bond now yields a higher return than the market interest rate, and we say it is trading at a premium.

**Example 6.5.4.** *The Orange Computer Company needs to raise money to expand. It issues a 10-year \$1,000 bond that pays \$30 every six months. If the current market interest rate is 7%, what is the fair market value of the bond?*

**Solution 6.5.4.** *The bond certificate promises us two things – An amount of \$1,000 to be paid in 10 years, and a semi-annual payment of \$30 for ten years. Therefore, to find the fair market value of the bond, we need to find the present value of the lump sum of \$1,000 we are to receive in 10 years, as well as, the present value of the \$30 semi-annual payments for the 10 years.*

*We will let  $P_1$  be the present value of the (face amount of \$1,000*

$$P_1 \left( 1 + \frac{0.07}{2} \right)^{20} = \$1,000$$

*Since the interest is paid twice a year, the interest is compounded twice a year and  $n = 2(10) = 20$*

$$P_1(1.9898) = \$1,000$$

$$P_1 = \$502.56$$

We will let  $P_2$  be the present value of the \$30 semi-annual payments is

$$P_2 \left(1 + \frac{0.07}{2}\right)^{20} = \$30 \left[ \frac{\left(1 + \frac{0.07}{2}\right)^{20} - 1}{\left(\frac{0.07}{2}\right)} \right]$$

$$P_2(1.9898) = 848.39$$

$$P_2 = \$426.37$$

The present value of the lump-sum \$1,000 = \$502.56 The present value of the \$30 semi-annual payments = \$426.37 The fair market value of the bond is  $P = P_1 + P_2 = \$502.56 + \$426.37 = \$928.93$

Note that because the market interest rate of 7% is higher than the bond's implied interest rate of 6% implied by the semiannual payments, the bond is selling at a discount; its fair market value of \$928.93 is less than its face value of \$1000.

**Example 6.5.5.** A state issues a 15 year \$1000 bond that pays \$25 every six months. If the current market interest rate is 4%, what is the fair market value of the bond?

**Solution 6.5.5.** The bond certificate promises two things – an amount of \$1,000 to be paid in 15 years, and semi-annual payments of \$25 for 15 years. To find the fair market value of the bond, we find the present value of the \$1,000 face value we are to receive in 15 years and add it to the present value of the \$25 semi-annual payments for the 15 years. In this example,  $n = 2(15) = 30$ .

We will let  $P_1$  be the present value of the lump-sum \$1,000

$$P_1(1 + 0.04/2)^{30} = \$1,000$$

$$P_1 = \$552.07$$

We will let  $P_2$  be the present value of the \$25 semi-annual payments is

$$P_2(1 + 0.04/2)^{30} = \$25 \left[ \frac{(1 + 0.04/2)^{30} - 1}{(0.04/2)} \right]$$

$$P_2(1.18114) = \$1014.20$$

$$P_2 = \$559.90$$

The present value of the lump-sum  $\$1,000 = \$552.07$  The present value of the  $\$25$  semi-annual payments  $= \$559.90$  Therefore, the fair market value of the bond is

$$P = P_1 + P_2 = \$552.07 + \$559.90 = \$1111.97$$

Because the market interest rate of 4% is lower than the interest rate of 5% implied by the semiannual payments, the bond is selling at a premium: the fair market value of  $\$1,111.97$  is more than the face value of  $\$1,000$ .

#### Summary 6.5.2: Fair Market Value of a Bond

To Find the Fair Market Value of a Bond, first find the present value of the face amount  $A$  that is payable at the maturity date:

$$A = P_1 \left(1 + \frac{r}{n}\right)^{nt}$$

Solve to find  $P_1$ .

Then find the present value of the semiannually payments of  $Sm$  over the term of the bond:

$$P_2 \left(1 + \frac{r}{n}\right)^{nt} = \frac{m \left[\left(1 + \frac{r}{n}\right)^{nt} - 1\right]}{r/n}$$

Solve to find  $P_2$ . The fair market value (or present value or price or current value) of the bond is the sum of the present values calculated above:

$$P = P_1 + P_2$$

#### 6.5.4 Amortization Schedule for a Loan

An amortization schedule is a table that lists all payments on a loan, splits them into the portion devoted to interest and the portion that is applied to repay principal, and calculates the outstanding balance on the loan after each payment is made.

**Example 6.5.6.** An amount of  $\$500$  is borrowed for 6 months at a rate of 12%. Make an amortization schedule showing the monthly payment, the

monthly interest on the outstanding balance, the portion of the payment contributing toward reducing the debt, and the outstanding balance.

**Solution 6.5.6.** The reader can verify that the monthly payment is \$86.27.

The first month, the outstanding balance is \$500, and therefore, the monthly interest on the outstanding balance is

$$(\text{outstanding balance})(\text{monthly interest rate}) = (\$500)(0.12/12) = \$5$$

This means, the first month, out of the \$86.27 payment, \$5 goes toward the interest and the remaining \$81.27 toward the balance leaving a new balance of  $\$500 - \$81.27 = \$418.73$ .

Similarly, the second month, the outstanding balance is \$418.73, and the monthly interest on the outstanding balance is \$4.19. Again, out of the \$86.27 payment, \$4.19 goes toward the interest and the remaining \$82.08 toward the balance leaving a new balance of  $\$418.73 - \$82.08 = \$336.65$ . The process continues in the table below.

Payment #	Payment	Interest	Debt Payment	Balance
1	\$86.27	\$5	\$81.27	\$418.73
2	\$86.27	\$4.19	\$82.08	\$336.65
3	\$86.27	\$3.37	\$82.90	\$253.75
4	\$86.27	\$2.54	\$83.73	\$170.02
5	\$86.27	\$1.70	\$84.57	\$85.45
6	\$86.27	\$0.85	\$85.42	\$0.03

Note that the last balance of 3 cents is due to error in rounding off.

An amortization schedule is usually lengthy and tedious to calculate by hand. For example, an amortization schedule for a 30-year mortgage loan with monthly payments would have  $(12)(30) = 360$  rows of calculations in the amortization schedule table. A car loan with 5 years of monthly payments would have  $12(5) = 60$  rows of calculations in the amortization schedule table. However, it would be straightforward to use a spreadsheet application on a computer or write a little code to do these repetitive calculations.

Most of the other applications in this section's problem set are reasonably straightforward and can be solved by taking a little extra care in interpreting them. And remember, there is often more than one way to solve a problem.



## 6.6 Classification of Finance Problems

In this section, you will review the concepts of Chapter 6 to:

1. re-examine the types of financial problems and classify them.
2. re-examine the vocabulary words used in describing financial calculations

We'd like to remind the reader that the hardest part of solving a finance problem is determining the category it falls into. So in this section, we will emphasize the classification of problems rather than finding the actual solution. We suggest that the student read each problem carefully and look for the word or words that may give clues to the kind of problem that is presented. For instance, students often fail to distinguish a lump-sum problem from an annuity. Since the payments are made each period, an annuity problem contains words such as each, every, per, etc. One should also be aware that in the case of a lump-sum, only a single deposit is made, while in an annuity numerous deposits are made at equal spaced time intervals. To help interpret the vocabulary used in the problems, we include a glossary at the end of this section.

Students often confuse the present value with the future value. For example, if a car costs \$15,000, then this is its present value. Surely, you cannot convince the dealer to accept \$15,000 in some future time, say, in five years. Recall how we found the installment payment for that car. We assumed that two people, Mr. Cash and Mr. Credit, were buying two identical cars both costing \$15,000 each. To settle the argument that both people should pay exactly the same amount, we put Mr. Cash's cash of \$15,000 in the bank as a lump-sum and Mr. Credit's monthly payments of  $x$  dollars each as an annuity. Then we make sure that the future values of these two accounts are equal. As you remember, at an interest rate of 9%

The future value of Mr. Cash's lump-sum was

$$\$15,000 \left( 1 + \frac{0.09}{12} \right)^{60}.$$

The future value of Mr. Credit's annuity was

$$\frac{m \left[ \left( 1 + \frac{0.09}{12} \right)^{60} - 1 \right]}{\frac{0.09}{12}}.$$

To solve the problem, we set the two expressions equal and solve for  $m$ . The present value of an annuity is found in exactly the same way. For example, suppose Mr. Credit is told that he can buy a particular car for \$311.38 a month for five years, and Mr. Cash wants to know how much he needs to pay. We are finding the present value of the annuity of \$311.38 per month, which is the same as finding the price of the car. This time our unknown quantity is the price of the car. Now suppose the price of the car is  $P$ , then the future value of Mr. Cash's lump-sum is

$$P \left( 1 + \frac{0.09}{12} \right)^{60},$$

and the future value of Mr. Credit's annuity is

$$\frac{\$311.38 \left[ \left( 1 + \frac{0.09}{12} \right)^{60} - 1 \right]}{\frac{0.09}{12}}.$$

Setting them equal we get,

$$P \left( 1 + \frac{0.09}{12} \right)^{60} = \frac{\$311.38 \left[ \left( 1 + \frac{0.09}{12} \right)^{60} - 1 \right]}{\frac{0.09}{12}}$$

$$P(1.5657) = (\$311.38)(75.4241)$$

$$P(1.5657) = \$23,485.57$$

$$P = \$15,000.04$$

### 6.6.1 Classification of Problems and Equations for Solutions

We now list six problems that form a basis for all finance problems. Further, we classify these problems and give an equation for the solution.

**Example 6.6.1.** *If \$2,000 is invested at 7% compounded quarterly, what will the final amount be in 5 years?*

**Solution 6.6.1.** *This is about the future (accumulated) value of a lump-sum. The future value can be calculated using the formula for compound interest:*

$$FV = A = \$2000 \left(1 + \frac{0.07}{4}\right)^{4 \cdot 5}$$

**Example 6.6.2.** *How much should be invested at 8% compounded yearly, for the final amount to be \$5,000 in five years?*

**Solution 6.6.2.** *This is about the present value of a lump-sum. The present value required for the future value of \$5,000 can be calculated as:*

$$PV(1 + 0.08)^5 = \$5,000$$

**Example 6.6.3.** *If \$200 is invested each month at 8.5% compounded monthly, what will the final amount be in 4 years?*

**Solution 6.6.3.** *The future value of an annuity formula is used in this case:*

$$FV = A = \frac{\$200 \left[ \left(1 + \frac{0.085}{12}\right)^{12 \times 4} - 1 \right]}{\frac{0.085}{12}}$$

**Example 6.6.4.** *How much should be invested each month at 9% for it to accumulate to \$8,000 in three years?*

**Solution 6.6.4.** *This is a sinking fund payment problem where the formula is:*

$$m \frac{\left[ \left(1 + \frac{0.09}{12}\right)^{12 \cdot 3} - 1 \right]}{\frac{0.09}{12}} = \$8,000$$

**Example 6.6.5.** *Keith has won a lottery paying him \$2,000 per month for the next 10 years. He'd rather have the entire sum now. If the interest rate is 7.6%, how much should he receive?*

**Solution 6.6.5.** *This is about the present value of an annuity. The present value of an annuity is calculated with:*

$$PV \left( 1 + \frac{0.076}{12} \right)^{12 \cdot 10} = \frac{\$2,000 \left[ \left( 1 + \frac{0.076}{12} \right)^{12 \cdot 10} - 1 \right]}{\frac{0.076}{12}}$$

**Example 6.6.6.** *Mr. A has just donated \$25,000 to his alma mater. Mr. B would like to donate an equivalent amount, but would like to pay by monthly payments over a five year period. If the interest rate is 8.2%, determine the size of the monthly payment?*

**Solution 6.6.6.** *The monthly payment can be found using the formula for the present value of an annuity due to the installment payment plan:*

$$m \frac{\left[ \left( 1 + \frac{0.082}{12} \right)^{60} - 1 \right]}{\frac{0.082}{12}} = \$25,000 \left( 1 + \frac{0.082}{12} \right)^{60}$$

## 6.7 Financial Calculations Vocabulary

As we've seen in these examples, it's important to read the problems carefully to correctly identify the situation. It is essential to understand the vocabulary for financial problems. Many of the vocabulary words used are listed in the glossary below for easy reference.

**Definition 6.7.1.** *The **Term**, denoted by  $t$ , is the time period for a loan or investment. In this book  $t$  is represented in years and should be converted into years when it is stated in months or other units.*

**Definition 6.7.2.** *The **Principal**, denoted by  $P$ , refers to the amount of money borrowed in a loan. If a sum of money is invested for a period of time, the sum invested at the start is the Principal.*

**Definition 6.7.3.** *The **Present Value**, denoted by  $P$ , is the value of money at the beginning of the time period.*

**Definition 6.7.4.** *The **Accumulated Value** or **Future Value** refers to the value of money at the end of the time period.*

**Definition 6.7.5.** *The **Discount** occurs in loans involving simple interest if the interest is deducted from the loan amount at the beginning of the loan period, rather than being repaid at the end of the loan period.*

**Definition 6.7.6.** The **Periodic Payment**, denoted by  $m$ , is the amount of a constant periodic payment that occurs at regular intervals during the time period under consideration, such as periodic payments made to repay a loan, regular periodic payments into a bank account as savings, or regular periodic payments to a retired person as an annuity.

**Definition 6.7.7.** The **Number of payment periods and compounding periods per year**, denoted by  $n$ , is considered to be the same in this book when dealing with periodic payments. While in general the compounding and payment periods do not have to be the same and calculations can become more complicated, formulas for different periods can be found in finance textbooks or various online resources. Technology such as online financial calculators, spreadsheet financial functions, or financial pocket calculators can be utilized for these calculations.

**Definition 6.7.8.** The **Number of periods**, denoted by  $nt$ , is calculated as  $nt = (\text{number of periods per year})(\text{number of years})$ . It gives the total number of payment and compounding periods.

**Definition 6.7.9.** The **Annual interest rate** or **Nominal rate** is the stated annual interest rate. This is expressed as a percent but converted to decimal form when used in financial calculation formulas. For example, if a bank account pays 3% interest compounded quarterly, then 3% is the nominal rate, and it is included in the financial formulas as  $r = 0.03$ .

**Definition 6.7.10.** The **Interest rate per compounding period**,  $\frac{r}{n}$ , is the interest rate for each compounding period. If a bank account pays 3% interest compounded quarterly, then  $\frac{r}{n} = \frac{0.03}{4} = 0.0075$ , corresponding to a rate of 0.75% per quarter. Some Finite Math books use the symbol  $i$  to represent  $\frac{r}{n}$ .

**Definition 6.7.11.** The **Effective Rate**, denoted by  $r_{\text{effective}}$ , or **Effective Annual Interest Rate**, also known as **Annual Percentage Yield (APY)** or **Annual Percentage Rate (APR)**, is the interest rate compounded annually that would yield the same interest as the stated compounded rate for the investment. The effective rate provides a uniform way for investors or borrowers to compare different interest rates with different compounding periods.

**Definition 6.7.12.** **Interest**, denoted by  $I$ , is money paid by a borrower for the use of money borrowed as a loan. It is also money earned over time when

depositing money into a savings account, certificate of deposit, or money market account. When a person deposits money in a bank account, the depositor is essentially lending money to the bank temporarily, and the bank pays interest to the depositor.

**Definition 6.7.13.** A **Sinking Fund** is a fund established by making periodic payments into a savings or investment account over a period of time. The purpose of a sinking fund is to save for a future purchase, such as a business setting aside money to buy equipment at the end of the savings period.

**Definition 6.7.14.** An **Annuity** is a series of periodic payments. In this book, it refers to a stream of constant periodic payments made at the end of each compounding period for a certain amount of time. Commonly, the term annuity is used to describe a steady stream of payments received by an individual as retirement income, like from a pension. Annuity payments may be made at the end of each payment period (ordinary annuity) or at the beginning (annuity due). While compounding and payment periods can differ, this textbook only addresses cases where these periods are the same.

**Definition 6.7.15.** A **Lump Sum** refers to a single sum of money paid or deposited all at once, rather than distributed over time. An example includes lottery winnings when the recipient opts for a one-time "lump sum" payment instead of periodic payments over time. The term "lump sum" implies that the transaction is a one-off and not a sequence of periodic payments.

**Definition 6.7.16.** A **Loan** is an amount of money borrowed with an agreement that the borrower will repay the lender in the future, within a specified period known as the term of the loan. Repayment typically occurs through periodic payments until the loan is fully paid off by the end of the term. Some loans may be repaid in a single sum at the loan's end, with interest paid either periodically during the term or as a lump sum at the end, or through a discount at the start of the loan.

# Chapter 7

## Sets and Counting

In this chapter, you will learn to:

1. Use set theory and Venn diagrams to solve counting problems.
2. Use the Multiplication Axiom to solve counting problems.
3. Use Permutations to solve counting problems.
4. Use Combinations to solve counting problems.
5. Use the Binomial Theorem to expand  $(x + y)^n$ .

### 7.1 Sets

In this section, you will learn to:

1. Use set notation to represent unions, intersections, and complements of sets.
2. Use Venn diagrams to solve counting problems.

#### 7.1.1 Introduction to Sets

In this section, we will familiarize ourselves with set operations and notations, so that we can apply these concepts to both counting and probability problems. We begin by defining some terms.

A set is a collection of objects, and its members are called the elements of the set. We name the set by using capital letters and enclose its members in braces. Suppose we need to list the members of the chess club. We use the following set notation:

$$C = \{\text{Ken, Bob, Tran, Shanti, Eric}\}$$

### 7.1.2 Empty Set, Set Equality, Subsets

A set that has no members is called an empty set. The empty set is denoted by the symbol  $\emptyset$ .

Two sets are equal if they have the same elements.

A set  $A$  is a subset of a set  $B$  if every member of  $A$  is also a member of  $B$ . Suppose  $C = \{\text{Al, Bob, Chris, David, Ed}\}$  and  $A = \{\text{Bob, David}\}$ . Then  $A$  is a subset of  $C$ , written as  $A \subseteq C$ .

Every set is a subset of itself, and the empty set is a subset of every set.

### 7.1.3 Union of Two Sets

Let  $A$  and  $B$  be two sets, then the union of  $A$  and  $B$ , written as  $A \cup B$ , is the set of all elements that are either in  $A$  or in  $B$ , or in both  $A$  and  $B$ .

### 7.1.4 Intersection of Two Sets

Let  $A$  and  $B$  be two sets, then the intersection of  $A$  and  $B$ , written as  $A \cap B$ , is the set of all elements that are common to both sets  $A$  and  $B$ .

A universal set  $U$  is the set consisting of all elements under consideration.

### 7.1.5 Complement of a Set, Disjoint Sets

Let  $A$  be any set, then the complement of set  $A$ , written as  $\overline{A}$ , is the set consisting of elements in the universal set  $U$  that are not in  $A$ . This can also be written as  $U - A$ , which is said “ $U$  less  $A$ ”.



Two sets  $A$  and  $B$  are called disjoint sets if their intersection is an empty set. Clearly, a set and its complement are disjoint; however, two sets can be disjoint and not be complements.

**Example 7.1.1.** List all the subsets of the set of primary colors  $\{\text{red}, \text{yellow}, \text{blue}\}$ .

**Solution 7.1.1.** The subsets are  $\emptyset$ ,  $\{\text{red}\}$ ,  $\{\text{yellow}\}$ ,  $\{\text{blue}\}$ ,  $\{\text{red}, \text{yellow}\}$ ,  $\{\text{red}, \text{blue}\}$ ,  $\{\text{yellow}, \text{blue}\}$ ,  $\{\text{red}, \text{yellow}, \text{blue}\}$ .

Note that the empty set is a subset of every set, and a set is a subset of itself.

**Example 7.1.2.** Let  $F = \{\text{Aikman}, \text{Jackson}, \text{Rice}, \text{Sanders}, \text{Young}\}$ , and let  $B = \{\text{Griffey}, \text{Jackson}, \text{Sanders}, \text{Thomas}\}$ . Find the intersection of the sets  $F$  and  $B$ .

**Solution 7.1.2.** The intersection of the two sets is the set whose elements belong to both sets. Therefore,  $F \cap B = \{\text{Jackson}, \text{Sanders}\}$

**Example 7.1.3.** Find the union of the sets  $F$  and  $B$  given as follows.

$$F = \{\text{Aikman}, \text{Jackson}, \text{Rice}, \text{Sanders}, \text{Young}\}$$

$$B = \{\text{Griffey}, \text{Jackson}, \text{Sanders}, \text{Thomas}\}$$

**Solution 7.1.3.** The union of two sets is the set whose elements are either in  $A$  or in  $B$  or in both  $A$  and  $B$ . Observe that when writing the union of two sets, the repetitions are avoided.

$$F \cup B = \{\text{Aikman}, \text{Griffey}, \text{Jackson}, \text{Rice}, \text{Sanders}, \text{Thomas}, \text{Young}\}$$

**Example 7.1.4.** Let the universal set  $U = \{\text{red}, \text{orange}, \text{yellow}, \text{green}, \text{blue}, \text{indigo}, \text{violet}\}$  and  $P = \{\text{red}, \text{yellow}, \text{blue}\}$ . Find the complement of  $P$ .

**Solution 7.1.4.** The complement of a set  $P$  is the set consisting of elements in the universal set  $U$  that are not in  $P$ . Therefore,

$$\overline{P} = \{\text{orange}, \text{green}, \text{indigo}, \text{violet}\}$$

To achieve a better understanding, suppose that the universal set  $U$  represents the colors of the spectrum, and  $P$  represents those colors of the spectrum that are not primary colors.

**Example 7.1.5.** Let the universal set  $U = \{\text{red, orange, yellow, green, blue, indigo, violet}\}$  and  $P = \{\text{red, yellow, blue}\}$ . Find a set  $R$  so that  $R$  is not the complement of  $P$  but  $R$  and  $P$  are disjoint.

**Solution 7.1.5.**  $R = \{\text{orange, green}\}$  and  $P = \{\text{red, yellow, blue}\}$  are disjoint because the intersection of the two sets is the empty set. The sets have no elements in common. However, they are not complements because their union  $P \cup R = \{\text{red, yellow, blue, orange, green}\}$  is not equal to the universal set  $U$ .

**Example 7.1.6.** Let  $U = \{\text{red, orange, yellow, green, blue, indigo, violet}\}$ ,  $P = \{\text{red, yellow, blue}\}$ ,  $Q = \{\text{red, green}\}$ , and  $R = \{\text{orange, green, indigo}\}$ . Find  $P \cup Q \cap \overline{R}$ .

**Solution 7.1.6.** We do the problems in steps:

$$\begin{aligned} P \cup Q &= \{\text{red, yellow, blue, green}\} \\ \overline{P \cup Q} &= \{\text{orange, indigo, violet}\} \\ \overline{R} &= \{\text{red, yellow, blue, violet}\} \\ \overline{P \cup Q} \cap \overline{R} &= \{\text{violet}\} \end{aligned}$$

## 7.1.6 Venn Diagrams

We now use Venn diagrams to illustrate the relationships between sets. In the late 1800s, an English logician named John Venn developed a method to represent relationships between sets. He represented these relationships using diagrams, which are now known as Venn diagrams.

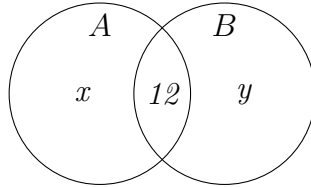
A Venn diagram represents a set as the interior of a circle. Often two or more circles are enclosed in a rectangle where the rectangle represents the universal set. To visualize an intersection or union of a set is easy. In this section, we will mainly use Venn diagrams to sort various populations and count objects.

**Example 7.1.7.** Suppose a survey of car enthusiasts showed that over a certain time period, 30 drove cars with automatic transmissions, 20 drove cars with standard transmissions, and 12 drove cars of both types. If everyone in the survey drove cars with one of these transmissions, how many people participated in the survey?

**Solution 7.1.7.** We will use Venn diagrams to solve this problem.

Let the set  $A$  represent those car enthusiasts who drove cars with automatic transmissions, and set  $S$  represent the car enthusiasts who drove the cars with standard transmissions. Now we use Venn diagrams to sort out the information given in this problem.

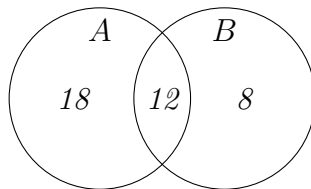
Since 12 people drove both cars, we place the number 12 in the region common to both sets.  $x$  represents the number of people who only drove cars with automatic transmissions. Similarly,  $y$  represents the number of drivers who only drove cars with standard transmissions.



Because 30 people drove cars with automatic transmissions, the circle  $A$  must contain 30 elements. This means that  $x + 12 = 30$ , or  $x = 18$ .

Similarly, since 20 people drove cars with standard transmissions, the circle  $S$  must contain 20 elements.

Thus,  $y + 12 = 20$  which in turn makes  $y = 8$ .

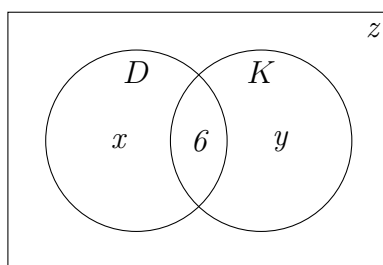


Now that all the information is sorted out, it is easy to read from the diagram that 18 people drove cars with automatic transmissions only, 12 people drove both types of cars, and 8 drove cars with standard transmissions only.

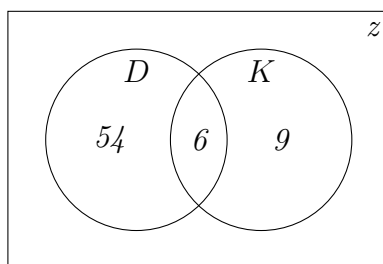
Therefore,  $18 + 12 + 8 = 38$  people took part in the survey.

**Example 7.1.8.** A survey of 100 people in California indicates that 60 people have visited Disneyland, 15 have visited Knott's Berry Farm, and 6 have visited both. How many people have visited neither place?

**Solution 7.1.8.** The problem is similar to the one in Example 7.1.7. Let the set  $D$  represent the people who have visited Disneyland, and  $K$  the set of people who have visited Knott's Berry Farm.

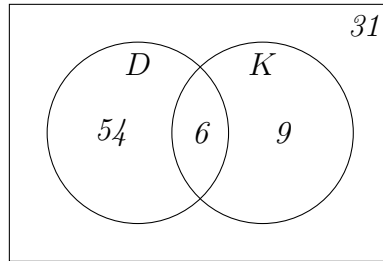


We fill the three regions associated with the sets  $D$  and  $K$  in the same manner as before. The number  $x$  represents the number of people that have been to Disneyland, but not Knotts. We know the total number of people who have been to Disneyland is 60, so  $x + 6 = 60$ . This means  $x$  must be 54. Similar reasoning will help us find that  $y = 9$ .



Since 100 people participated in the survey, the rectangle representing the universal set  $U$  must contain 100 objects. Let  $z$  represent those people in the universal set that are neither in the set  $D$  nor in  $K$ . This means  $54 + 6 + 9 + z = 100$ , or  $z = 31$ .

Therefore, there are 31 people in the survey who have visited neither place.



**Example 7.1.9.** A survey of 100 exercise-conscious people resulted in the following information:

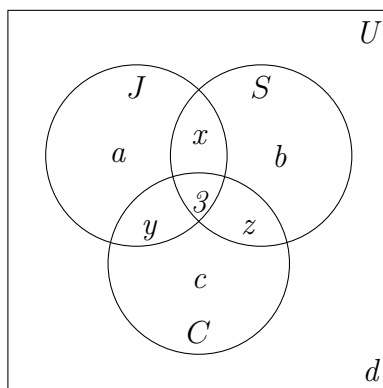
- 50 jog, 30 swim, and 35 cycle
- 14 jog and swim
- 7 swim and cycle
- 9 jog and cycle
- 3 people take part in all three activities

1. How many jog but do not swim or cycle?
2. How many take part in only one of the activities?
3. How many do not take part in any of these activities?

**Solution 7.1.9.** Let  $J$  represent the set of people who jog,  $S$  the set of people who swim, and  $C$  who cycle.

In using Venn diagrams, our ultimate aim is to assign a number to each region. We always begin by first assigning the number to the innermost region and then working our way out.

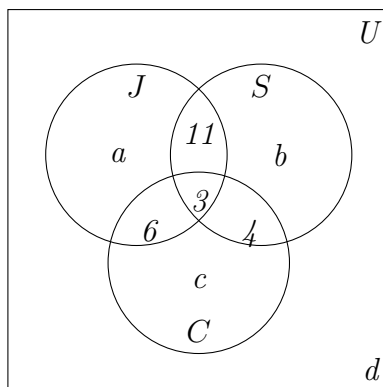
We'll show the solution step by step. As you practice working out such problems, you will find that with practice you will not need to draw multiple copies of the diagram.



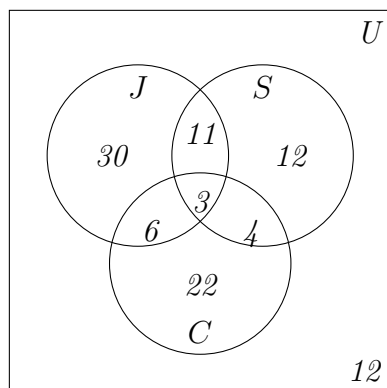
We place a 3 in the innermost region because it represents the number of people who participate in all three activities. Next, we label the regions we don't know about. Now, we compute  $x$ ,  $y$ , and  $z$ .

Since 14 people jog and swim,  $x + 3 = 14$ , or  $x = 11$ . The fact that 9 people jog and cycle results in  $y + 3 = 9$ , or  $y = 6$ . Since 7 people swim and cycle,  $z + 3 = 7$ , or  $z = 4$ .

Updating our Venn diagram



Now we proceed to find the unknowns  $a$ ,  $b$ , and  $c$ , as shown in Figure IV. Since 50 people jog,  $a + 11 + 6 + 3 = 50$ , or  $a = 30$ . 30 people swim, therefore,  $b + 11 + 4 + 3 = 30$ , or  $b = 12$ . 35 people cycle, therefore,  $c + 6 + 4 + 3 = 35$ , or  $c = 22$ . By adding all the entries in all three sets, we get a sum of 88. Since 100 people were surveyed, the number,  $d$ , inside the universal set but outside of all three sets is given by  $100 - 88 = 12$ . Finally we put all the information in our diagram, all the information is sorted out, and the questions can readily be answered.



1. The number of people who jog but do not swim or cycle is 30.
2. The number who take part in only one of these activities is  $30 + 12 + 22 = 64$ .
3. The number of people who do not take part in any of these activities is 12.

## 7.2 Tree Diagrams and the Multiplication Axiom

In this section you will learn to

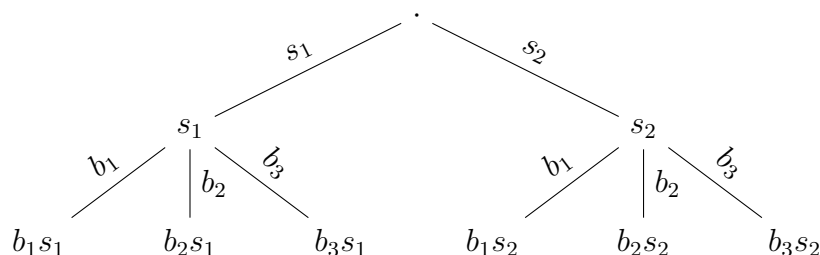
1. Use trees to count possible outcomes in a multi-step process
2. Use the multiplication axiom to count possible outcomes in a multi-stop process.

In this chapter, we are trying to develop counting techniques that will be used in the next chapter to study probability. One of the most fundamental of such techniques is called the Multiplication Axiom. Before we introduce the multiplication axiom, we first look at some examples.

**Example 7.2.1.** *If a woman has two blouses and three skirts, how many different outfits consisting of a blouse and a skirt can she wear?*

**Solution 7.2.1.** *Suppose we call the blouses  $b_1$  and  $b_2$ , and skirts  $s_1$ ,  $s_2$ , and  $s_3$ . We can have the following six outfits:  $b_1s_1$ ,  $b_1s_2$ ,  $b_1s_3$ ,  $b_2s_1$ ,  $b_2s_2$ ,  $b_2s_3$*

*Alternatively, we can draw a tree diagram:*



The tree diagram gives us all six possibilities. The method involves two steps. First the woman chooses a blouse. She has two choices: blouse one or blouse two. If she chooses blouse one, she has three skirts to match it with; skirt one, skirt two, or skirt three. Similarly if she chooses blouse two, she can match it with each of the three skirts, again. The tree diagram helps us visualize these possibilities.

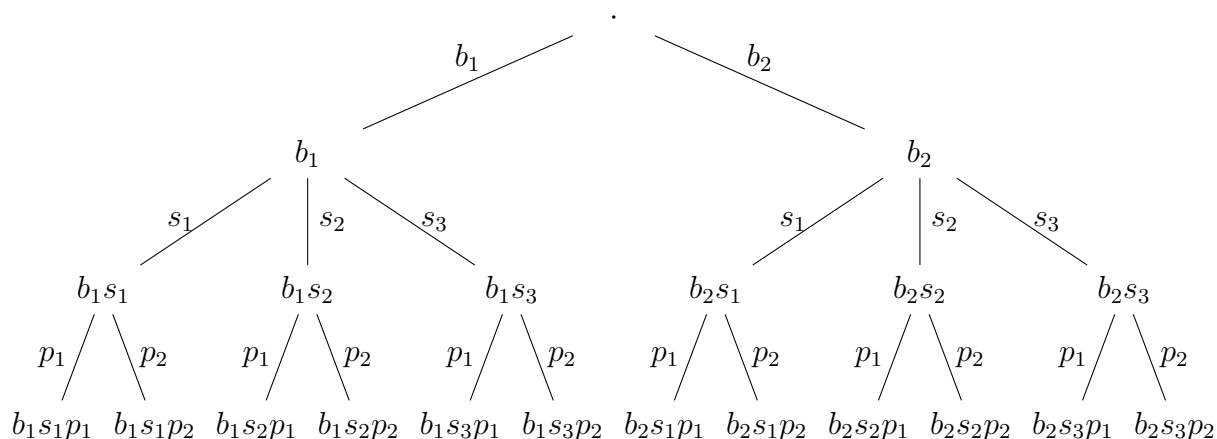
The reader should note that the process involves two steps. For the first step of choosing a blouse, there are two choices, and for each choice of a blouse, there are three choices of choosing a skirt. So altogether there are  $2 \cdot 3 = 6$  possibilities.

If, in the previous example, we add the shoes to the outfit, we have the following problem.

**Example 7.2.2.** *If a woman has two blouses, three skirts, and two pumps, how many different outfits consisting of a blouse, a skirt, and a pair of pumps can she wear?*

**Solution 7.2.2.** *Suppose we call the blouses  $b_1$  and  $b_2$ , the skirts  $s_1$ ,  $s_2$ , and  $s_3$ , and the pumps  $p_1$ , and  $p_2$ . The following tree diagram will illustrate all the possible outfits.*





We count the number of branches in the tree, and see that there are 12 different possibilities.

This time the method involves three steps. First, the woman chooses a blouse. She has two choices: blouse one or blouse two. Now suppose she chooses blouse one. This takes us to step two of the process which consists of choosing a skirt. She has three choices for a skirt, and let us suppose she chooses skirt two. Now that she has chosen a blouse and a skirt, we have moved to the third step of choosing a pair of pumps. Since she has two pairs of pumps, she has two choices for the last step. Let us suppose she chooses pumps two. She has chosen the outfit consisting of blouse one, skirt two, and pumps two, or  $b_1s_2p_2$ . By looking at the different branches on the tree, one can easily see the other possibilities.

The important thing to observe here, again, is that this is a three step process. There are two choices for the first step of choosing a blouse. For each choice of a blouse, there are three choices of choosing a skirt, and for each combination of a blouse and a skirt, there are two choices of selecting a pair of pumps.

All in all, we have  $2 \cdot 3 \cdot 2 = 12$  different possibilities.

Tree diagrams help us visualize the different possibilities, but they are not practical when the possibilities are numerous. Besides, we are mostly interested in finding the number of elements in the set and not the actual list of all possibilities; once the problem is envisioned, we can solve it without a tree diagram. The two examples we just solved may have given us a clue to

do just that.

Let us now try to solve Example 2 without a tree diagram. The problem involves three steps: choosing a blouse, choosing a skirt, and choosing a pair of pumps. The number of ways of choosing each are listed below. By multiplying these three numbers we get 12, which is what we got when we did the problem using a tree diagram.

The number of ways to choose a blouse	The number of ways to choose a skirt	The number of ways to choose pumps
2	3	2

The procedure we just employed is called the multiplication axiom.

#### Summary 7.2.1: THE MULTIPLICATION AXIOM

If a task can be done in  $m$  ways, and a second task can be done in  $n$  ways, then the operation involving the first task followed by the second can be performed in  $m \cdot n$  ways.

The general multiplication axiom is not limited to just two tasks and can be used for any number of tasks.

**Example 7.2.3.** *A truck license plate consists of a letter followed by four digits. How many such license plates are possible?*

**Solution 7.2.3.** *Since there are 26 letters and 10 digits, we have the following choices for each.*

Letter	Digit	Digit	Digit	Digit
26	10	10	10	10

Therefore, the number of possible license plates is  $26 \cdot 10 \cdot 10 \cdot 10 \cdot 10 = 260000$ .

**Example 7.2.4.** *In how many different ways can a 3-question true-false test be answered?*

**Solution 7.2.4.** *Since there are two choices for each question, we have*

Question 1	Question 2	Question 3
2	2	2

Applying the multiplication axiom, we get  $2 \cdot 2 \cdot 2 = 8$  different ways.

We list all eight possibilities:  $TTT$ ,  $TTF$ ,  $TFT$ ,  $TFF$ ,  $FTT$ ,  $FTF$ ,  $FFT$ ,  $FFF$ .

The reader should note that the first letter in each possibility is the answer corresponding to the first question, the second letter corresponds to the answer to the second question, and so on. For example,  $TFF$ , says that the answer to the first question is given as true, and the answers to the second and third questions false.

**Example 7.2.5.** *In how many different ways can four people be seated in a row?*

**Solution 7.2.5.** *Suppose we put four chairs in a row, and proceed to put four people in these seats. There are four choices for the first chair we choose. Once a person sits down in that chair, there are only three choices for the second chair, and so on. We list as shown below.*

$$\boxed{4} \mid \boxed{3} \mid \boxed{2} \mid \boxed{1}$$

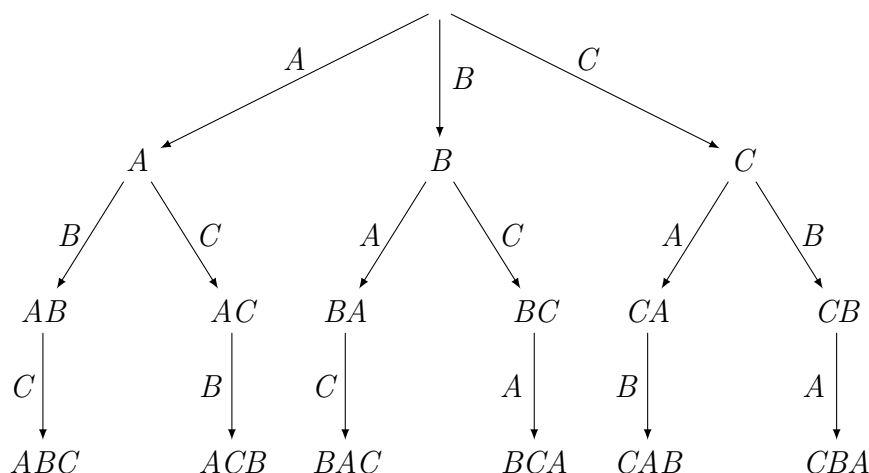
*So there are altogether  $4 \cdot 3 \cdot 2 \cdot 1 = 24$  different ways.*

**Example 7.2.6.** *How many three-letter word sequences can be formed using the letters  $\{A, B, C\}$  if no letter is to be repeated?*

**Solution 7.2.6.** *The problem is very similar to the previous example. Imagine a child having three building blocks labeled  $A$ ,  $B$ , and  $C$ . Suppose he puts these blocks on top of each other to make word sequences. For the first letter he has three choices, namely  $A$ ,  $B$ , or  $C$ . Let us suppose he chooses the first letter to be a  $B$ , then for the second block which must go on top of the first, he has only two choices:  $A$  or  $C$ . And for the last letter he has only one choice. We list the choices below.*

$$\boxed{3} \mid \boxed{2} \mid \boxed{1}$$

*Therefore,  $3 \cdot 2 \cdot 1 = 6$  different word sequences can be formed.*



### 7.2.1 Permutations

In this section, you will learn to:

1. Count the number of possible permutations (ordered arrangement) of  $n$  items taken  $r$  at a time.
2. Count the number of possible permutations when there are conditions imposed on the arrangements.
3. Perform calculations using factorials.

In Example 7.2.6, we were asked to find the word sequences formed by using the letters  $\{A, B, C\}$  if no letter is to be repeated. The tree diagram gave us the following six arrangements:  $ABC$ ,  $ACB$ ,  $BAC$ ,  $BCA$ ,  $CAB$ , and  $CBA$ .

Arrangements like these, where order is important and no element is repeated, are called permutations.

**Definition 7.2.1.** A **permutation** of a set of elements is an ordered arrangement where each element is used once.

**Example 7.2.7.** How many three-letter word sequences can be formed using the letters  $\{A, B, C, D\}$ ?

**Solution 7.2.7.** There are four choices for the first letter of our word, three choices for the second letter, and two choices for the third.

$$\boxed{4 \mid 3 \mid 2}$$

Applying the multiplication axiom, we get  $4 \cdot 3 \cdot 2 = 24$  different arrangements.

**Example 7.2.8.** How many permutations of the letters of the word *ARTICLE* have consonants in the first and last positions?

**Solution 7.2.8.** In the word *ARTICLE*, there are 4 consonants.

Since the first letter must be a consonant, we have four choices for the first position, and once we use up a consonant, there are only three consonants left for the last spot. We show as follows:

$$\begin{array}{|c|c|c|c|c|c|c|} \hline 4 & ? & ? & ? & ? & ? & 3 \\ \hline \end{array}$$

Since there are no more restrictions, we can go ahead and make the choices for the rest of the positions.

So far we have used up 2 letters, therefore, five remain. So for the next position there are five choices, for the position after that there are four choices, and so on. We get

$$\begin{array}{|c|c|c|c|c|c|c|} \hline 4 & 5 & 4 & 3 & 2 & 1 & 3 \\ \hline \end{array}$$

So the total permutations are  $4 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 \cdot 3 = 1440$ .

**Example 7.2.9.** Given five letters  $\{A, B, C, D, E\}$ . Find the following:

- The number of four-letter word sequences.
- The number of three-letter word sequences.
- The number of two-letter word sequences.

**Solution 7.2.9.** The problem is easily solved by the multiplication axiom, and answers are as follows:

- The number of four-letter word sequences is  $5 \cdot 4 \cdot 3 \cdot 2 = 120$ .
- The number of three-letter word sequences is  $5 \cdot 4 \cdot 3 = 60$ .
- The number of two-letter word sequences is  $5 \cdot 4 = 20$ .

We often encounter situations where we have a set of  $n$  objects and we are selecting  $r$  objects to form permutations. We refer to this as permutations of  $n$  objects taken  $r$  at a time, and we write it as  $nPr$ .

Therefore, the above example can also be answered as listed below:

1. The number of four-letter word sequences is  $5P4 = 120$ .
2. The number of three-letter word sequences is  $5P3 = 60$ .
3. The number of two-letter word sequences is  $5P2 = 20$ .

Before we give a formula for  $nPr$ , we'd like to introduce a symbol that we will use a great deal in this as well as in the next chapter.

**Definition 7.2.2.** The **factorial** of a natural number  $n$  is given by

$$n! = n \cdot (n - 1) \cdot (n - 2) \cdot \dots \cdot 3 \cdot 2 \cdot 1$$

Also, we define  $0! = 1$ .

**Definition 7.2.3.** The number of permutations of  $n$  objects taken  $r$  at a time (with  $n$  and  $r$  natural numbers) denoted  $nPr$  is given by either

$$nPr = n \cdot (n - 1) \cdot (n - 2) \cdot \dots \cdot (n - r + 1)$$

or

$$nPr = \frac{n!}{(n - r)!}$$

**Example 7.2.10.** Compute the following using both formulas.

1.  $6P3$
2.  $7P2$

**Solution 7.2.10.** We will identify  $n$  and  $r$  in each case and solve using the formulas provided.

1.  $6P3 = 6 \cdot 5 \cdot 4 = 120$ , or

$$6P3 = \frac{6!}{(6 - 3)!} = \frac{6!}{3!} = \frac{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1} = 120$$

2.  $7P2 = 7 \cdot 6 = 42$ , or

$$7P2 = \frac{7!}{5!} = \frac{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 42$$

Next we consider some more permutation problems to get further insight into these concepts.

**Example 7.2.11.** *In how many different ways can 4 people be seated in a straight line if two of them insist on sitting next to each other?*

**Solution 7.2.11.** *Let us suppose we have four people A, B, C, and D. Further suppose that A and B want to sit together. For the sake of argument, we glue A and B together and treat them as one person  $\boxed{AB}$ .*

*The four people are  $\boxed{AB}$  CD. Since  $\boxed{AB}$  is treated as one person, we have the following possible arrangements.*

$$\boxed{AB} CD, \boxed{AB} DC, C \boxed{AB} D, D \boxed{AB} C, CD \boxed{AB}, DC \boxed{AB}$$

*Note that there are six more such permutations because A and B could also be glued together in the order  $\boxed{BA}$ . And they are*

$$\boxed{BA} CD, \boxed{BA} DC, C \boxed{BA} D, D \boxed{BA} C, CD \boxed{BA}, DC \boxed{BA}$$

*So altogether there are 12 different permutations.*

*Let us now do the problem using the multiplication axiom.*

*After we glue two of the people together and treat them as one person, we can say we have only three people. The multiplication axiom tells us that three people can be seated in  $3!$  ways. Since two people can be glued together  $2!$  ways, there are  $3! \times 2! = 12$  different arrangements*

**Example 7.2.12.** *You have 4 math books and 5 history books to put on a shelf that has 5 slots. In how many ways can the books be shelved if the first three slots are filled with math books and the next two slots are filled with history books?*

**Solution 7.2.12.** *We first do the problem using the multiplication axiom.*

*Since the math books go in the first three slots, there are 4 choices for the first slot, 3 choices for the second and 2 choices for the third.*

*The fourth slot requires a history book, and has five choices. Once that choice is made, there are 4 history books left, and therefore, 4 choices for the last slot. The choices are shown below.*

4	3	2	5	4
---	---	---	---	---

Therefore, the number of permutations are  $4 \cdot 3 \cdot 2 \cdot 5 \cdot 4 = 480$ .

Alternately, we can see that  $4 \cdot 3 \cdot 2$  is really same as  $4P3$ , and  $5 \cdot 4$  is  $5P2$ .

So the answer can be written as  $(4P3)(5P2) = 480$ .

Clearly, this makes sense. For every permutation of three math books placed in the first three slots, there are  $5P2$  permutations of history books that can be placed in the last two slots. Hence the multiplication axiom applies, and we have the answer  $(4P3)(5P2)$ .

## 7.3 Circular Permutations and Permutations with Similar Elements

In this section you will learn to:

1. Count the number of possible permutations of items arranged in a circle.
2. Count the number of possible permutations when there are repeated items.

In this section, we will address the following two problems:

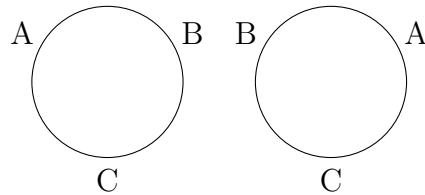
1. In how many different ways can five people be seated in a circle?
2. In how many different ways can the letters of the word MISSISSIPPI be arranged?

The first problem comes under the category of Circular Permutations, and the second under Permutations with Similar Elements.

### 7.3.1 Circular Permutations

Suppose we have three people named A, B, and C. We have already determined that they can be seated in a straight line in  $3!$  or 6 ways. Our next problem is to see how many ways these people can be seated in a circle. We draw a diagram.





It happens that there are only two ways we can seat three people in a circle, relative to each other's positions. This kind of permutation is called a circular permutation. In such cases, no matter where the first person sits, the permutation is not affected. Each person can shift as many places as they like, and the permutation will not be changed. We are interested in the position of each person in relation to the others. Imagine the people on a merry-go-round; the rotation of the permutation does not generate a new permutation. So in circular permutations, the first person is considered a place holder, and where they sit does not matter.

#### Summary 7.3.1: Circular Permutations

The number of permutations of  $n$  elements in a circle is  $(n - 1)!$ .

Another way to look at this is to say that there are  $n!$  ways to place the people around the circle, but we have to divide by  $n$  because there are  $n$  different people to consider the start.

**Example 7.3.1.** *In how many different ways can five people be seated at a circular table?*

**Solution 7.3.1.** *We have already determined that the first person is just a place holder. Therefore, there is only one choice for the first spot. We have*

$$\boxed{1} \mid \boxed{4} \mid \boxed{3} \mid \boxed{2} \mid \boxed{1}$$

*So the answer is 24.*

**Example 7.3.2.** *At a scholarship dinner, there are four professors and four students. How many ways are there to sit if they want to sit alternately?*

**Solution 7.3.2.** *We again emphasize that the first person can sit anywhere without affecting the permutation.*

*So there is only one choice for the first spot. Suppose a student sat down*

first (since they're generally hungrier). The chair to the right must belong to a professor, and there are 4 choices. The next chair belongs to a student, so there are three choices and so on. We list the choices below.

1	4	3	3	2	2	1	1
---	---	---	---	---	---	---	---

So the answer is 144.

### 7.3.2 Permutations with Similar Elements

Let us determine the number of distinguishable permutations of the letters ELEMENT. Suppose we make all the letters different by labeling the letters as follows.

$E_1LE_2ME_3NT$ .

Since all the letters are now different, there are  $7!$  different permutations.

Let us now look at one such permutation, say

$LE_1ME_2NE_3T$

Suppose we form new permutations from this arrangement by only moving the E's. Clearly, there are  $3!$  or 6 such arrangements. We list them below.

$LE_1ME_2NE_3T$   
 $LE_1ME_3NE_2T$   
 $LE_2ME_1NE_3T$   
 $LE_2ME_3NE_1T$   
 $LE_3ME_1NE_2T$   
 $LE_3ME_2NE_1T$

Because the E's are not different, there is only one arrangement LEMENET and not six. This is true for every permutation.

Let us suppose there are  $n$  different permutations of the letters ELEMENT.

Then there are  $n \cdot 3!$  permutations of the letters  $E_1LE_2ME_3NT$ .

### 7.3. CIRCULAR PERMUTATIONS AND PERMUTATIONS WITH SIMILAR ELEMENTS 211

But we know there are  $7!$  permutations of the letters  $E_1LE_2ME_3NT$ .

Therefore,  $n \cdot 3! = 7!$

Or  $n = \frac{7!}{3!}$ .

#### Summary 7.3.2: Permutations with Similar Elements

The number of permutations of  $n$  elements taken  $n$  at a time, with  $r_1$  elements of one kind,  $r_2$  elements of another kind, and so on, is given by

$$\frac{n!}{r_1! r_2! \dots r_k!}$$

**Example 7.3.3.** Find the number of different permutations of the letters of the word MISSISSIPPI.

**Solution 7.3.3.** The word MISSISSIPPI has 11 letters. If the letters were all different there would have been  $11!$  different permutations. But MISSISSIPPI has 4 S's, 4 I's, and 2 P's that are alike.

So the answer is

$$\frac{11!}{4!4!2!} = 34,650.$$

**Example 7.3.4.** If a coin is tossed six times, how many different outcomes consisting of 4 heads and 2 tails are there?

**Solution 7.3.4.** Again, we have permutations with similar elements.

We are looking for permutations for the letters HHHHTT.

The answer is

$$\frac{6!}{4!2!} = 15.$$

**Example 7.3.5.** In how many different ways can 4 nickels, 3 dimes, and 2 quarters be arranged in a row?

**Solution 7.3.5.** Assuming that all nickels are similar, all dimes are similar, and all quarters are similar, we have permutations with similar elements. Therefore, the answer is

$$\frac{9!}{4!3!2!} = 1260.$$

**Example 7.3.6.** *A stock broker wants to assign 20 new clients equally to 4 of its salespeople. In how many different ways can this be done?*

**Solution 7.3.6.** *This means that each salesperson gets 5 clients. The problem can be thought of as an ordered partitions problem. In that case, using the formula we get*

$$\frac{20!}{5!5!5!5!} = 11,732,745,024.$$

**Example 7.3.7.** *A college has a straight row of 5 flagpoles at its main entrance. It has 3 identical green flags and 2 identical yellow flags. How many distinct arrangements of flags on the flagpoles are possible?*

**Solution 7.3.7.** *The problem can be thought of as distinct permutations of the letters GGGYY; that is arrangements of 5 letters, where 3 letters are similar, and the remaining 2 letters are similar:*

$$\frac{5!}{3!2!} = 10$$

*Just to provide a little more insight into the solution, we list all 10 distinct permutations: GGGYY, GGYGY, GGYYG, GYGGY, GYGYG, GYYGG, YGGGY, YGGYG, YGYGG, YYGGG*

## 7.4 Combinations

In this section, you will learn to:

1. Count the number of combinations of  $r$  out of  $n$  items (selections without regard to arrangement).
2. Use factorials to perform calculations involving combinations.

Suppose we have a set of three letters {A, B, C}, and we are asked to make two-letter word sequences. We have the following six permutations:

$$AB \quad BA \quad BC \quad CB \quad AC \quad CA$$

Now suppose we have a group of three people {A, B, C} as Alex, Blake, and Cam, respectively, and we are asked to form committees of two people each.

This time we have only three committees, namely:

$$AB \quad BC \quad AC$$

When forming committees, the order is not important because the committee that has Alex and Blake is no different than the committee that has Blake and Alex. As a result, we have only three committees and not six.

Forming word sequences is an example of permutations, while forming committees is an example of combinations - the topic of this section.

Permutations are those arrangements where order is important, while combinations are those arrangements where order is not significant. From now on, this is how we will tell permutations and combinations apart.

In the above example, there were six permutations, but only three combinations.

Just as the symbol  $nPr$  represents the number of permutations of  $n$  objects taken  $r$  at a time,  $nCr$  represents the number of combinations of  $n$  objects taken  $r$  at a time.

So in the above example,  $3P2 = 6$ , and  $3C2 = 3$ .

Our next goal is to determine the relationship between the number of combinations and the number of permutations in a given situation.

In the above example, if we knew that there were three combinations, we could have found the number of permutations by multiplying this number by  $2!$ . That is because each combination consists of two letters, and that makes  $2!$  permutations.

**Example 7.4.1.** *Given the set of letters  $\{A, B, C, D\}$ . Write the number of combinations of three letters, and then from these combinations determine the number of permutations.*

**Solution 7.4.1.** *We have the following four combinations.*

$$ABC, BCD, CDA, BDA$$

*Since every combination has three letters, there are  $3!$  permutations for every combination. We list them below.*

$ABC$	$ACB$	$BAC$	$BCA$	$CAB$	$CBA$
$BCD$	$BDC$	$CBD$	$CDB$	$DBC$	$DCB$
$CDA$	$CAD$	$DCA$	$DAC$	$ACD$	$ADC$
$BDA$	$BAD$	$DBA$	$DAB$	$ABD$	$ADB$

The number of permutations are  $3!$  times the number of combinations; that is

$${}_4P_3 = 3! \cdot {}_4C_3$$

or

$${}_4C_3 = \frac{{}_4P_3}{3!}$$

In general,

$${}_nC_r = \frac{{}_nP_r}{r!}$$

Since

$${}_nP_r = \frac{n!}{(n-r)!}$$

We have,

$${}_nC_r = \frac{n!}{(n-r)! \cdot r!}$$

#### Summary 7.4.1: Combinations

A combination of a set of elements is an arrangement where each element is used once, and order is not important. The number of combinations of  $n$  objects taken  $r$  at a time is given by

$${}_nC_r = \frac{n!}{(n-r)! \cdot r!}$$

where  $n$  and  $r$  are natural numbers.

**Example 7.4.2.** *Compute:*

1.  ${}_5C_3$
2.  ${}_7C_3$

**Solution 7.4.2.** *We use the above formula.*

$$1. {}^5C_3 = \frac{5!}{(5-3)!3!} = \frac{5!}{2!3!} = 10$$

$$2. {}^7C_3 = \frac{7!}{(7-3)!3!} = \frac{7!}{4!3!} = 35$$

**Example 7.4.3.** *In how many different ways can a student select to answer five questions from a test that has seven questions, if the order of the selection is not important?*

**Solution 7.4.3.** *Since the order is not important, it is a combination problem, and the answer is*

$${}^7C_5 = 21.$$

**Example 7.4.4.** *How many line segments can be drawn by connecting any two of the six points that lie on the circumference of a circle?*

**Solution 7.4.4.** *Since the line that goes from point A to point B is same as the one that goes from B to A, this is a combination problem. It is a combination of 6 objects taken 2 at a time. Therefore, the answer is*

$${}^6C_2 = \frac{6!}{4!2!} = 15.$$

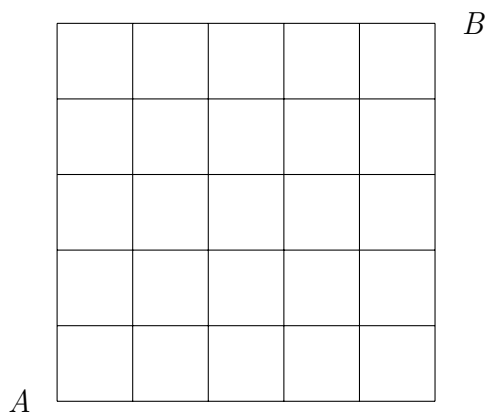
**Example 7.4.5.** *There are ten people at a party. If they all shake hands, how many hand-shakes are possible?*

**Solution 7.4.5.** *Note that between any two people there is only one hand shake. Therefore, we have*

$${}^{10}C_2 = 45 \text{ hand-shakes.}$$

**Example 7.4.6.** *The shopping area of a town is in the shape of square that is 5 blocks by 5 blocks. How many different routes can a taxi driver take to go from one corner of the shopping area to the opposite cater-corner?*

**Solution 7.4.6.** *Let us suppose the taxi driver drives from the point A, the lower left hand corner, to the point B, the upper right hand corner as shown in the figure below.*



To reach his destination, he has to travel ten blocks; five horizontal, and five vertical. So if out of the ten blocks he chooses any five horizontal, the other five will have to be the vertical blocks, and vice versa.

Therefore, all he has to do is to choose 5 out of ten to be the horizontal blocks

The answer is  $10C5$ , or 252.

Alternately, the problem can be solved by permutations with similar elements. The taxi driver's route consists of five horizontal and five vertical blocks. If we call a horizontal block  $H$ , and a vertical block  $V$ , then one possible route may be as follows.

$HHHHHVVVVV$

Clearly there are

$$\frac{10!}{5! \cdot 5!} = 252$$

permutations.

Further note that by definition

$$10C5 = \frac{10!}{5! \cdot 5!}.$$

**Example 7.4.7.** If a coin is tossed six times, in how many ways can it fall four heads and two tails?



**Solution 7.4.7.** *First, we solve this problem using techniques from 7.3.2 - permutations with similar elements.*

*We need 4 heads and 2 tails, that is we need to find all the rearrangements of HHHHTT. Since there are 4 Hs and 2 Ts. The number of permutations is given by*

$$\frac{6!}{4!2!} = 15.$$

*Now we solve this problem using combinations.*

*Suppose we have six spots to put the coins on. If we choose any four spots for heads, the other two will automatically be tails. So the problem is simply*

$${}^6C_4 = 15.$$

*Incidentally, we could have easily chosen the two tails, instead. In that case, we would have gotten*

$${}^6C_2 = 15.$$

*Further observe that by definition*

$${}^6C_4 = \frac{6!}{2!4!} \text{ and } {}^6C_2 = \frac{6!}{4!2!}$$

**Note 7.4.1.** *It is true that,*

$${}^nC_r = \frac{n!}{r!(n-r)!} = \frac{n!}{(n-r)!(r)!} = {}^nC_{(n-r)}$$

## 7.5 Combinations Involving Several Sets

In this section, you will learn to:

1. Count the number of items selected from more than one set.
2. Count the number of items selected when there are restrictions on the selections.

So far, we have solved the basic combination problem of  $r$  objects chosen from  $n$  different objects. Now we will consider certain variations of this problem.

**Example 7.5.1.** *How many five-person committees consisting of 2 faculty members and 3 students can be chosen from a total of 4 faculty members and 4 students?*

**Solution 7.5.1.** *We list 4 faculty members and 4 students as follows:*

$$F_1 F_2 F_3 F_4 S_1 S_2 S_3 S_4$$

*Since we want 5-person committees consisting of 2 faculty members and 3 students, we'll first form all possible two-faculty committees and all possible three-student committees. Clearly there are  $4C2 = 6$  two-faculty committees, and  $4C3 = 4$  three-student committees, we list them as follows:*

*2-Faculty Committees*

- $F_1 F_2$
- $F_1 F_3$
- $F_1 F_4$
- $F_2 F_3$
- $F_2 F_4$
- $F_3 F_4$

*3-Student Committees*

- $S_1 S_2 S_3$
- $S_1 S_2 S_4$
- $S_1 S_3 S_4$
- $S_2 S_3 S_4$

*For every 2-faculty committee there are four 3-student committees that can be chosen to make a 5-person committee. If we choose  $F_1 F_2$  as our 2-faculty committee, then we can choose any of  $S_1 S_2 S_3$ ,  $S_1 S_2 S_4$ ,  $S_1 S_3 S_4$ , or  $S_2 S_3 S_4$  as our 3-student committees. As a result, we get*

$$\boxed{F_1F_2}S_1S_2S_3, \boxed{F_1F_2}S_1S_2S_4, \boxed{F_1F_2}S_1S_3S_4, \boxed{F_1F_2}S_2S_3S_4$$

Similarly, if we choose  $F_1F_3$  as our 2-faculty committee, then, again, we can choose any of  $S_1S_2S_3$ ,  $S_1S_2S_4$ ,  $S_1S_3S_4$ , or  $S_2S_3S_4$  as our 3-student committees.

$$\boxed{F_1F_3}S_1S_2S_3, \boxed{F_1F_3}S_1S_2S_4, \boxed{F_1F_3}S_1S_3S_4, \boxed{F_1F_3}S_2S_3S_4$$

And so on.

Since there are six 2-faculty committees, and for every 2-faculty committee there are four 3-student committees, there are altogether  $6 \cdot 4 = 24$  five-person committees.

In essence, we are applying the multiplication axiom to the different combinations.

**Example 7.5.2.** A club consists of 4 freshmen, 5 sophomores, 5 juniors, and 6 seniors. How many ways can a committee of 4 people be chosen that includes:

1. One student from each class?
2. All juniors?
3. Two freshmen and 2 seniors?
4. No freshmen?
5. At least three seniors?

**Solution 7.5.2.** 1. Applying the multiplication axiom to the combinations involved, we get

$$(4C1)(5C1)(5C1)(6C1) = 600$$

2. We are choosing all 4 members from the 5 juniors, and none from the others.

$$5C4 = 5$$

3.  $4C2 \cdot 6C2 = 90$

4. Since we don't want any freshmen on the committee, we need to choose all members from the remaining 16. That is

$${}_{16}C_4 = 1820$$

5. Of the 4 people on the committee, we want at least three seniors. This can be done in two ways. We could have three seniors, and one non-senior, or all four seniors.

$$({}_6C_3)({}_{14}C_1) + {}_6C_4 = 295$$

**Example 7.5.3.** How many five-letter word sequences consisting of 2 vowels and 3 consonants can be formed from the letters of the word *INTRODUCE*?

**Solution 7.5.3.** First we select a group of five letters consisting of 2 vowels and 3 consonants. Since there are 4 vowels and 5 consonants, we have

$$({}_4C_2)({}_5C_3)$$

Since our next task is to make word sequences out of these letters, we multiply these by 5!:

$$({}_4C_2)({}_5C_3)(5!) = 7200.$$

### 7.5.1 A Standard Deck of 52 Playing Cards

As in the previous example, many examples and homework problems in this book refer to a standard deck of 52 playing cards. Before we end this section, we take a minute to describe a standard deck of playing cards, as some readers may not be familiar with this.

A standard deck of 52 playing cards has 4 suits with 13 cards in each suit. The suits are diamonds ( $\heartsuit$ ), hearts ( $\heartsuit$ ), spades ( $\spadesuit$ ), and clubs ( $\clubsuit$ ). Each suit is associated with a color, either black ( $\spadesuit$ ,  $\clubsuit$ ) or red ( $\heartsuit$ ,  $\diamondsuit$ ).

Each suit contains 13 denominations (or values) for cards: the nine numbers 2, 3, 4, ..., 10 and Jack(J), Queen (Q), King (K), Ace (A).

The Jack, Queen and King are called "face cards" because they have pictures on them. Therefore a standard deck has 12 face cards: (3 values JQK)  $\times$  (4 suits  $\heartsuit \diamondsuit \spadesuit \clubsuit$ ). There are two Jacks in profile ( $\heartsuit$ ,  $\spadesuit$ ) and two Jacks in full

face ( $\diamond$ ,  $\clubsuit$ ), the Jacks in profile are sometimes referred to as the one-eyed Jacks.

We can visualize the 52 cards by the following display

Suit	Color	Values (Denominations)
$\diamond$ Diamonds	Red	2 3 4 5 6 7 8 9 10 J Q K A
$\heartsuit$ Hearts	Red	2 3 4 5 6 7 8 9 10 J Q K A
$\spadesuit$ Spades	Black	2 3 4 5 6 7 8 9 10 J Q K A
$\clubsuit$ Clubs	Black	2 3 4 5 6 7 8 9 10 J Q K A

**Example 7.5.4.** *A standard deck of playing cards has 52 cards consisting of 4 suits with 13 cards in each. In how many different ways can a 5-card hand consisting of four cards of one suit and one of another be drawn?*

**Solution 7.5.4.** *We will do the problem using the following steps.*

1. *Select a suit.*
2. *Select four cards from this suit.*
3. *Select another suit.*
4. *Select a card from that suit.*

*Applying the multiplication axiom, we have*

$$\begin{array}{ll}
 \text{Ways of selecting the first suit} & 4C1 \\
 \text{Ways of selecting 4 cards from this suit} & 13C4 \\
 \text{Ways of selecting the next suit} & 3C1 \\
 \text{Ways of selecting a card from that suit} & 13C1 \\
 \hline
 (4C1) \cdot (13C4) \cdot (3C1) \cdot (13C1) & = 11,540.
 \end{array}$$

## 7.6 Binomial Theorem

In this section, you will learn to:

1. find the coefficients of a binomial expansion such as  $(x + y)^n$  quickly and accurately

We end this chapter with one more application of combinations. Combinations are used in determining the coefficients of a binomial expansion such as  $(x + y)^n$ . Expanding a binomial expression by multiplying it out is a very



We now replace the blanks with the coefficients in equation (I), and we get

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

**Solution 7.6.1.** The expansion  $(x + y)^7$  is  $(x + y)(x + y)(x + y)(x + y)(x + y)(x + y)(x + y)$ . In multiplying the right side, each product is gotten by picking an  $x$  or  $y$  from each of the seven factors  $(x + y)$ . The term  $x^2y^5$  is obtained by choosing an  $x$  from two of the factors and a  $y$  from the other five factors. This can be done in  ${}^7C_2$ , or 21 ways. Therefore, the coefficient of the term  $x^2y^5$  is 21.

**Solution 7.6.2.** We first write the expansion without the coefficients.

$$(x+y)^7 = \underline{\hspace{1cm}}x^7 + \underline{\hspace{1cm}}x^6y + \underline{\hspace{1cm}}x^5y^2 + \underline{\hspace{1cm}}x^4y^3 + \underline{\hspace{1cm}}x^3y^4 + \underline{\hspace{1cm}}x^2y^5 + \underline{\hspace{1cm}}xy^6 + \underline{\hspace{1cm}}y^7$$

- The coefficient of the term  $x^7$  is 7C7 or 7C0 which equals 1.
- The coefficient of the term  $x^6y$  is 7C6 or 7C1 which equals 7.
- The coefficient of the term  $x^5y^2$  is 7C5 or 7C2 which equals 21.
- The coefficient of the term  $x^4y^3$  is 7C4 or 7C3 which equals 35,
- and so on.

$$(x+y)^7 = x^7 + 7x^6y + 21x^5y^2 + 35x^4y^3 + 35x^3y^4 + 21x^2y^5 + 7xy^6 + y^7$$

**Summary 7.6.1: Binomial Theorem**

$$(x+y)^n = \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \cdots + \binom{n}{n-1}xy^{n-1} + \binom{n}{n}y^n$$

**Example 7.6.3.** Expand  $(3a - 2b)^4$ .

**Solution 7.6.3.** If we let  $x = 3a$  and  $y = -2b$ , and apply the Binomial Theorem, we get

$$\begin{aligned} & (3a - 2b)^4 \\ &= 4C0(3a)^4 + 4C1(3a)^3(-2b) + 4C2(3a)^2(-2b)^2 + 4C3(3a)(-2b)^3 + 4C4(-2b)^4 \\ &= 1(81a^4) + 4(27a^3)(-2b) + 6(9a^2)(4b^2) + 4(3a)(-8b^3) + 1(16b^4) \\ &= 81a^4 - 216a^3b + 216a^2b^2 - 96a^3b + 16b^4 \end{aligned}$$

**Example 7.6.4.** Find the fifth term of the expansion  $(3a - 2b)^7$ .

**Solution 7.6.4.** The Binomial theorem tells us that in the  $r$ -th term of an expansion, the exponent of the  $y$  term is always one less than  $r$ , and the coefficient of the term is  $nC_{r-1}$ . Thus, for  $n = 7$  and  $r - 1 = 5 - 1 = 4$ , so the coefficient is  $7C4 = 35$ . Therefore, the fifth term is

$$(7C4)(3a)^{7-4}(-2b)^4 = 35(27a^3)(16b^4) = 15120a^3b^4$$



# Chapter 8

## Probability

In this chapter, you will learn to:

1. Write sample spaces.
2. Determine whether two events are mutually exclusive.
3. Use the Addition Rule.
4. Calculate probabilities using both tree diagrams and combinations.
5. Solve problems involving conditional probability.
6. Determine whether two events are independent.

### 8.1 Sample Spaces and Probability

In this section, you will learn to:

1. Write sample spaces.
2. Calculate probabilities by examining simple events in sample spaces.

If two coins are tossed, what is the probability that both coins will land heads? The problem seems simple enough, but it is not uncommon to hear the incorrect answer of  $1/3$ . A student may incorrectly reason that if two coins are tossed, there are three possibilities: one head, two heads, or no heads. Therefore, the probability of two heads is one out of three. The

answer is wrong because if we toss two coins, there are four possibilities and not three. For clarity, assume that one coin is a penny and the other is a nickel. Then we have the following four possibilities:

- HH (Both coins land heads)
- HT (Penny lands heads, nickel lands tails)
- TH (Penny lands tails, nickel lands heads)
- TT (Both coins land tails)

### 8.1.1 Sample Spaces

An act of flipping coins, rolling dice, drawing cards, or surveying people are referred to as a probability experiment.

**Definition 8.1.1.** *The **sample space** of an experiment is the set of all possible outcomes.*

**Example 8.1.1.** *If a die is rolled, write a sample space.*

**Solution 8.1.1.** *A die has six faces each having an equally likely chance of appearing. Therefore, the set of all possible outcomes  $S$  is*

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

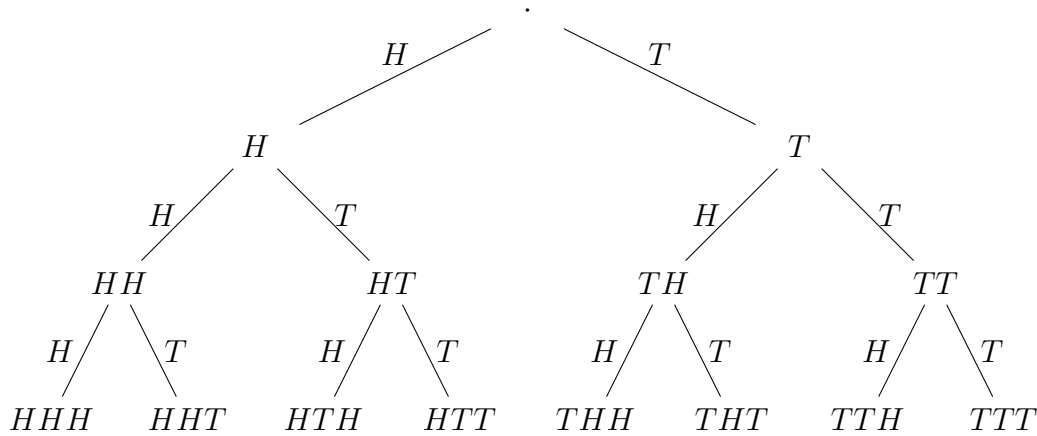
**Example 8.1.2.** *A coin is flipped three times. Write a sample space.*

**Solution 8.1.2.** *The sample space consists of eight possibilities.*

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

*The possibility HTH, for example, indicates that the first flip is a head, the second flip is a tail, and the third is a head.*

*We illustrate these possibilities with a tree diagram.*



**Example 8.1.3.** Two dice are rolled. Write the sample space.

**Solution 8.1.3.** We assume one of the dice is red, and the other green. We have the following 36 possibilities.

		Green					
		1	2	3	4	5	6
Red	1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
	2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
	3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
	4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
	5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
	6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

The entry (2, 5), for example, indicates that the red die shows a 2, and the green a 5.

### Summary 8.1.1: Probability

For a sample space  $S$  and an outcome  $A$  of  $S$ , the following two properties are satisfied.

1. If  $A$  is an outcome of a sample space, then the probability of  $A$ , denoted by  $P(A)$ , is between 0 and 1, inclusive.

$$0 \leq P(A) \leq 1$$

2. The sum of the probabilities of all the outcomes in  $S$  equals 1.

The probability  $P(A)$  of an event  $A$  describes the chance or likelihood of that event occurring.

- If  $P(A) = 0$ , event  $A$  will almost surely not occur.
- If  $P(A) = 1$ , event  $A$  will almost surely occur.
- If  $P(A) = 0.5$ , then event  $A$  is equally likely to occur or not occur.

**Note 8.1.1.** *It seems strange to use the phrase **almost surely** in that sentence. It seems like an event with probability zero is just impossible, but almost surely has a precise mathematical meaning. It's a bit beyond this course to explain, but we're going to try anyways. Suppose that you're throwing a dart at a dart board. The probability of hitting any single point is zero, but you've probably seen a dart hit a dart board, so it does happen. If you tried to hit the exact center of the board we'd expect it to take infinite time, but it's not impossible. In practice, we'd have a lot of trouble actually locating the exact center of the board because the markings on our measuring devices aren't thin enough...*

For example, if we toss a fair coin that is equally likely to land on heads or tails, then  $P(\text{Head}) = 0.50$ . If the weather forecast says there is a 70% chance of rain today, then  $P(\text{Rain}) = 0.70$ , indicating it is more likely to rain than to not rain.

**Example 8.1.4.** *One 6 sided die is rolled once. Find the probability that the result is greater than 4.*

**Solution 8.1.4.** *The sample space consists of the following six possibilities in set  $S$ :  $S = \{1, 2, 3, 4, 5, 6\}$*

*Let  $E$  be the event that the number rolled is greater than four:  $E = \{5, 6\}$*

*Therefore, the probability of  $E$  is:*

$$P(E) = \frac{2}{6} \text{ or } \frac{1}{3}.$$

**Example 8.1.5.** *If two dice, one red and one green, are rolled, find the probability that the red die shows a 3 and the green shows a six.*

**Solution 8.1.5.** *Since two dice are rolled, there are 36 possibilities. The probability of each outcome, listed in Example 8.1.3, is equally likely.*

Since  $(3, 6)$  is one such outcome, the probability of obtaining  $(3, 6)$  is  $\frac{1}{36}$ .

The example we just considered consisted of only one outcome of the sample space. We are often interested in finding probabilities of several outcomes represented by an event. An event is a subset of a sample space. If an event consists of only one outcome, it is called a simple event.

**Example 8.1.6.** *If two dice are rolled, find the probability that the sum of the faces of the dice is 7.*

**Solution 8.1.6.** *Let  $E$  represent the event that the sum of the faces of two dice is 7.*

*The possible cases for the sum to be equal to 7 are:  $(1, 6)$ ,  $(2, 5)$ ,  $(3, 4)$ ,  $(4, 3)$ ,  $(5, 2)$ , and  $(6, 1)$ , so event  $E$  is*

$$E = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$$

	1	2	3	4	5	6
1	$(1, 1)$	$(1, 2)$	$(1, 3)$	$(1, 4)$	$(1, 5)$	$(1, 6)$
2	$(2, 1)$	$(2, 2)$	$(2, 3)$	$(2, 4)$	$(2, 5)$	$(2, 6)$
3	$(3, 1)$	$(3, 2)$	$(3, 3)$	$(3, 4)$	$(3, 5)$	$(3, 6)$
4	$(4, 1)$	$(4, 2)$	$(4, 3)$	$(4, 4)$	$(4, 5)$	$(4, 6)$
5	$(5, 1)$	$(5, 2)$	$(5, 3)$	$(5, 4)$	$(5, 5)$	$(5, 6)$
6	$(6, 1)$	$(6, 2)$	$(6, 3)$	$(6, 4)$	$(6, 5)$	$(6, 6)$

*The probability of the event  $E$  is*

$$P(E) = \frac{6}{36} \text{ or } \frac{1}{6}.$$

**Example 8.1.7.** *A jar contains 3 red, 4 white, and 3 blue marbles. If a marble is chosen at random, what is the probability that the marble is a red marble or a blue marble?*

**Solution 8.1.7.** *We assume the marbles are  $r_1, r_2, r_3$  for red,  $w_1, w_2, w_3, w_4$  for white, and  $b_1, b_2, b_3$  for blue. Let the event  $C$  represent that the marble is red or blue.*

*The sample space  $S$  is  $\{r_1, r_2, r_3, w_1, w_2, w_3, w_4, b_1, b_2, b_3\}$*

And the event  $C = \{ r_1, r_2, r_3, b_1, b_2, b_3 \}$

Therefore, the probability of  $C$ ,

$$P(C) = \frac{6}{10} \text{ or } \frac{3}{5}.$$

**Example 8.1.8.** A jar contains three marbles numbered 1, 2, and 3. If two marbles are drawn without replacement, what is the probability that the sum of the numbers is 5?

**Solution 8.1.8.** Note, the two marbles in this example are drawn consecutively without replacement. That means that after a marble is drawn it is not replaced in the jar, and therefore is no longer available to select on the second draw.

Since two marbles are drawn without replacement, the sample space consists of the following six possibilities.

$$S = \{(1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2)\}$$

Note that  $(1, 1)$ ,  $(2, 2)$  and  $(3, 3)$  are not listed in the sample space. These outcomes are not possible when drawing without replacement, because once the first marble is drawn but not replaced into the jar, that marble is not available in the jar to be selected again on the second draw.

Let the event  $E$  represent that the sum of the numbers is five. Then

$$E = \{(2, 3), (3, 2)\}$$

Therefore, the probability of  $E$  is

$$P(E) = \frac{2}{6} \text{ or } \frac{1}{3}.$$

**Example 8.1.9.** A jar contains three marbles numbered 1, 2, and 3. If two marbles are drawn without replacement, what is the probability that the sum of the numbers is at least 4?

**Solution 8.1.9.** The sample space, as in Example 8.1.8, consists of the following six possibilities.

$$S = \{(1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2)\}$$

Let the event  $F$  represent that the sum of the numbers is at least four. Then

$$F = \{(1, 3), (3, 1), (2, 3), (3, 2)\}$$

Therefore, the probability of  $F$  is

$$P(F) = \frac{4}{6} \text{ or } \frac{2}{3}.$$

**Example 8.1.10.** A jar contains three marbles numbered 1, 2, and 3. If two marbles are drawn with replacement, what is the probability that the sum of the numbers is 5?

**Solution 8.1.10.** When two marbles are drawn with replacement, the sample space consists of the following nine possibilities.

$$S = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3)\}$$

Let the event  $E$  represent that the sum of the numbers is five. Then

$$E = \{(2, 3), (3, 2)\}$$

Therefore, the probability of  $E$  is

$$P(E) = \frac{2}{9}$$

Note that in Example 8.1.10 when we selected marbles with replacement, the probability has changed from Example 8.1.8 where we selected marbles without replacement.

**Example 8.1.11.** A jar contains three marbles numbered 1, 2, and 3. If two marbles are drawn with replacement, what is the probability that the sum of the numbers is at least 4?

**Solution 8.1.11.** *The sample space when drawing with replacement consists of the following nine possibilities.*

$$S = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3)\}$$

*Let the event  $F$  represent that the sum of the numbers is at least four. Then*

$$F = \{(1, 3), (3, 1), (2, 3), (3, 2), (2, 2), (3, 3)\}$$

*Therefore, the probability of  $F$  is*

$$P(F) = \frac{6}{9} \text{ or } \frac{2}{3}.$$

Note that in Example 8.1.11 when we selected marbles with replacement, the probability is the same as in Example 8.1.9 where we selected marbles without replacement.

Thus sampling with or without replacement MAY change the probabilities, but may not, depending on the situation in the particular problem under consideration. We'll re-examine the concepts of sampling with and without replacement in Section 8.3.

## 8.2 Mutually Exclusive Events and the Addition Rule

In this section, you will learn to:

1. Define compound events using union, intersection, and complement.
2. Identify mutually exclusive events
3. Use the Addition Rule to calculate probability for unions of events.

In Chapter 7, we learned to find the union, intersection, and complement of a set. We will now use these set operations to describe events.

The **union** of two events  $E$  and  $F$ ,  $E \cup F$ , is the set of outcomes that are in  $E$  or in  $F$  or in both.



## 8.2. MUTUALLY EXCLUSIVE EVENTS AND THE ADDITION RULE 233

The **intersection** of two events  $E$  and  $F$ ,  $E \cap F$ , is the set of outcomes that are in both  $E$  and  $F$ .

The **complement** of an event  $E$ , denoted by  $\overline{E}$ , is the set of outcomes in the sample space  $S$  that are not in  $E$ .

### Summary 8.2.1: The Complement Rule

If  $E^c$  is the Complement of Event  $E$ , then

$$P(E^c) = 1 - P(E)$$

This follows from the fact that if the sample space has  $n$  elements and  $E$  has  $k$  elements, then  $\overline{E}$  has  $n - k$  elements. Therefore,

$$P(\overline{E}) = \frac{n - k}{n} = 1 - \frac{k}{n} = 1 - P(E).$$

Of particular interest to us are the events whose outcomes do not overlap. We call these events mutually exclusive.

Two events  $E$  and  $F$  are said to be **mutually exclusive** if they do not intersect:  $E \cap F = \emptyset$ .

Next we'll determine whether a given pair of events are mutually exclusive.

**Example 8.2.1.** *A card is drawn from a standard deck. Determine whether the pair of events given below is mutually exclusive.*

$$E = \{\text{The card drawn is an Ace}\}$$

$$F = \{\text{The card drawn is a heart}\}$$

**Solution 8.2.1.** *Clearly the ace of hearts belongs to both sets. That is*

$$E \cap F = \{A\heartsuit\} \neq \emptyset.$$

*Therefore, the events  $E$  and  $F$  are not mutually exclusive.*

**Example 8.2.2.** *Two dice are rolled. Determine whether the pair of events given below is mutually exclusive.*

$$G = \{\text{The sum of the faces is six}\}$$

$$H = \{\text{One die shows a four}\}$$

**Solution 8.2.2.** *For clarity, we list the elements of both sets.*

$$G = \{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}$$

$$H = \{(2, 4), (4, 2)\}$$

Clearly,  $G \cap H = \{(2, 4), (4, 2)\} \neq \emptyset$ . Therefore, the two sets are not mutually exclusive.

**Example 8.2.3.** *A committee consists of three members. Determine whether the following pair of events are mutually exclusive.*

$$M = \{\text{The committee has at least one faculty member}\}$$

$$N = \{\text{The committee consists only of students}\}$$

**Solution 8.2.3.** *Although the answer may be clear, we list both the sets.*

$$M = \{FFF, FFS, FSF, FSS, SFF, SFS, SSF\}$$

$$N = \{SSS\}$$

Clearly,

$$M \cap N = \emptyset$$

Therefore, events  $M$  and  $N$  are mutually exclusive.

We will now consider problems that involve the union of two events.

Given two events,  $E$  and  $F$ , then finding the probability of  $E \cup F$  is the same as finding the probability that  $E$  will happen, or  $F$  will happen, or both will happen.

**Example 8.2.4.** *If a die is rolled, what is the probability of obtaining an even number or a number greater than four?*

**Solution 8.2.4.** *Let  $E$  be the event that the number shown on the die is an even number, and let  $F$  be the event that the number shown is greater than four.*

*The sample space  $S$  is  $\{1, 2, 3, 4, 5, 6\}$ . The event  $E = \{2, 4, 6\}$ , and event  $F = \{5, 6\}$ .*

*We need to find  $P(E \cup F)$ .*

## 8.2. MUTUALLY EXCLUSIVE EVENTS AND THE ADDITION RULE 235

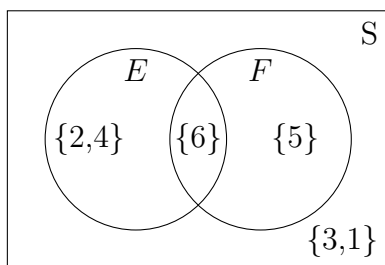
Since  $P(E) = \frac{3}{6}$ , and  $P(F) = \frac{2}{6}$ , a student may say  $P(E \cup F) = \frac{3}{6} + \frac{2}{6}$ . This will be incorrect because the element 6, which is in both  $E$  and  $F$  has been counted twice, once as an element of  $E$  and once as an element of  $F$ . In other words, the set  $E \cup F$  has only four elements and not five: set  $E \cup F = \{2, 4, 5, 6\}$

Therefore,  $P(E \cup F) = \frac{4}{6}$  and not  $\frac{5}{6}$ .

This can be illustrated by a Venn diagram. We'll use the Venn Diagram to re-examine Example 8.2.4 and derive a probability rule that we can use to calculate probabilities for unions of events.

The sample space  $S$ , the events  $E$  and  $F$ , and  $E \cap F$  are listed below.

$$S = \{1, 2, 3, 4, 5, 6\}, \quad E = \{2, 4, 6\}, \quad F = \{5, 6\}, \quad E \cap F = \{6\}.$$



Finding the probability of  $E \cup F$  is the same as finding the probability that  $E$  will happen, or  $F$  will happen, or both will happen.

If we count the number of elements  $\#(E)$  in  $E$ , and add to it the number of elements  $\#(F)$  in  $F$ , the points in both  $E$  and  $F$  are counted twice, once as elements of  $E$  and once as elements of  $F$ . Now if we subtract from the sum,  $\#(E) + \#(F)$ , the number  $\#(E \cap F)$ , we remove the duplicity and get the correct answer. So as a rule,

$$\#(E \cup F) = \#(E) + \#(F) - \#(E \cap F)$$

By dividing the entire equation by  $\#(S)$ , we get

$$\frac{\#(E \cup F)}{\#(S)} = \frac{\#(E)}{\#(S)} + \frac{\#(F)}{\#(S)} - \frac{\#(E \cap F)}{\#(S)}$$

Since the probability of an event is the number of elements in that event divided by the number of all possible outcomes, we have

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

Applying the above for Example 8.2.4, we get

$$P(E \cup F) = \frac{3}{6} + \frac{2}{6} - \frac{1}{6} = \frac{4}{6}$$

This is because, when we add  $P(E)$  and  $P(F)$ , we have added  $P(E \cap F)$  twice. Therefore, we must subtract  $P(E \cap F)$ , once.

This gives us the general formula, called the **Addition Rule**, for finding the probability of the union of two events. Because event  $E \cup F$  is the event that  $E$  will happen, OR  $F$  will happen, OR both will happen, we sometimes call this the **Addition Rule for OR Events**. It states

#### Summary 8.2.2: Addition Rule

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

If, and only if, two events  $E$  and  $F$  are mutually exclusive, then  $E \cap F = \emptyset$  and  $P(E \cap F) = 0$ , and we get

$$P(E \cup F) = P(E) + P(F)$$

**Example 8.2.5.** *If a card is drawn from a deck, use the addition rule to find the probability of obtaining an ace or a heart.*

**Solution 8.2.5.** *Let  $A$  be the event that the card is an ace, and  $H$  the event that it is a heart. Since there are four aces, and thirteen hearts in the deck,*

$$P(A) = \frac{4}{52} \text{ and } P(\heartsuit) = \frac{13}{52}.$$

*Furthermore, since the intersection of two events consists of only one card,  $A \cap \heartsuit$ , we now have:*

$$P(A \cap \heartsuit) = \frac{1}{52}$$

## 8.2. MUTUALLY EXCLUSIVE EVENTS AND THE ADDITION RULE 237

We need to find  $P(A \cup \heartsuit)$ :

$$\begin{aligned} P(A \cup \heartsuit) &= P(A) + P(\heartsuit) - P(A \cap \heartsuit) \\ &= \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52}. \end{aligned}$$

**Example 8.2.6.** Two dice are rolled, and the events  $F$  and  $T$  are as follows:  $F = \{\text{The sum of the dice is four}\}$  and  $T = \{\text{At least one die shows a three}\}$ . Find  $P(F \cup T)$ .

**Solution 8.2.6.** We list  $F$  and  $T$ , and  $F \cap T$  as follows:

$$F = \{(1, 3), (2, 2), (3, 1)\}$$

$$T = \{(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6), (1, 3), (2, 3), (4, 3), (5, 3), (6, 3)\}$$

$$F \cap T = \{(1, 3), (3, 1)\}$$

Since  $P(F \cup T) = P(F) + P(T) - P(F \cap T)$  We have  $P(F \cup T) = 3/36 + 11/36 - 2/36 = 12/36$ .

**Example 8.2.7.** Dr. T is seeking a mathematics instructor's position at the best community college in California, Columbia College. Their employment depends on two conditions: whether the board approves the position, and whether the hiring committee selects them. There is an 80% chance that the board will approve the position, and there is a 70% chance that the hiring committee will select them. If there is a 90% chance that at least one of the two conditions, the board approval or their selection, will be met, what is the probability that Dr. T will be hired?

**Solution 8.2.7.** Let  $A$  be the event that the board approves the position, and  $S$  be the event that Dr. T gets selected. We have,

$$P(A) = .80, \quad P(S) = .70, \quad P(A \cup S) = .90$$

We need to find,  $P(A \cap S)$ . The addition formula states that,

$$P(A \cup S) = P(A) + P(S) - P(A \cap S)$$

Substituting the known values, we get

$$.90 = .80 + .70 - P(A \cap S)$$

Therefore,  $P(A \cap S) = .60$ .

**Example 8.2.8.** *The probability that this weekend will be cold is 0.6, the probability that it will be rainy is 0.7, and the probability that it will be both cold and rainy is 0.5. What is the probability that it will be neither cold nor rainy?*

**Solution 8.2.8.** *Let  $C$  be the event that the weekend will be cold, and  $R$  be the event that it will be rainy. We are given that*

$$P(C) = 0.6, \quad P(R) = 0.7, \quad P(C \cap R) = 0.5$$

*First we find  $P(C \cup R)$  using the Addition Rule.*

$$P(C \cup R) = P(C) + P(R) - P(C \cap R) = 0.6 + 0.7 - 0.5 = 0.8$$

*Then we find  $P(\overline{C \cup R})$  using the Complement Rule.*

$$P(\overline{C \cup R}) = 1 - P(C \cup R) = 1 - 0.8 = 0.2$$

## 8.3 Probability Using Tree Diagrams and Combinations

In this section, you will learn to:

1. Use probability tree diagrams to calculate probabilities.
2. Use combinations to calculate probabilities.

In this section, we will apply previously learned counting techniques in calculating probabilities and use tree diagrams to help us gain a better understanding of what is involved.

### 8.3.1 Using Tree Diagrams to Calculate Probabilities

We already used tree diagrams to list events in a sample space. Tree diagrams can be helpful in organizing information in probability problems; they help provide a structure for understanding probability. In this section, we expand our previous use of tree diagrams to situations in which the events in the sample space are not all equally likely.

We assign the appropriate probabilities to the events shown on the branches of the tree. By multiplying probabilities along a path through the tree, we can find probabilities for "and" events, which are intersections of events.

We begin with an example.

**Example 8.3.1.** Suppose a jar contains 3 red and 4 white marbles. If two marbles are drawn with replacement, what is the probability that both marbles are red?

**Solution 8.3.1.** Let  $E$  be the event that the first marble drawn is red, and let  $F$  be the event that the second marble drawn is red.

We need to find  $P(E \cap F)$ . By the statement, "two marbles are drawn with replacement," we mean that the first marble is replaced before the second marble is drawn.

There are 7 choices for the first draw. Since the first marble is replaced before the second is drawn, there are again seven choices for the second draw. Using the multiplication axiom, we conclude that the sample space  $S$  consists of 49 ordered pairs. Of the 49 ordered pairs, there are  $3 \cdot 3 = 9$  ordered pairs that show red on the first draw and, also, red on the second draw. Therefore,

$$P(E \cap F) = \frac{9}{49}$$

Further note that in this particular case

$$P(E \cap F) = \frac{9}{49} = \frac{3}{7} \cdot \frac{3}{7}$$

giving us the result that in this example:  $P(E \cap F) = P(E) \cdot P(F)$

**Example 8.3.2.** If we use the same 3 red and 4 white marbles from Example 8.3.1, but the two marbles are drawn **without** replacement, what is the probability that both marbles are red?

**Solution 8.3.2.** By the statement, "two marbles are drawn without replacement," we mean that the first marble is not replaced before the second marble is drawn.

There are, again, 7 choices for the first draw but since the first marble is not replaced before the second is drawn, there are only six choices for the second draw. Using the multiplication axiom, we conclude that the sample space  $S$  consists of 42 ordered pairs. Of the 42 ordered pairs, there are  $3 \cdot 2 = 6$

ordered pairs that show red on the first draw and red on the second draw. Therefore,

$$P(E \cap F) = \frac{6}{42}$$

Note that we can break this calculation down as

$$P(E \cap F) = \frac{6}{42} = \frac{3}{7} \cdot \frac{2}{6}$$

Here  $\frac{3}{7}$  represents  $P(E)$ , and  $\frac{2}{6}$  represents the probability of drawing a red on the second draw, given that the first draw resulted in a red.

We write the latter as  $P(\text{red on the second} \mid \text{red on first})$  or  $P(F \mid E)$ . The “ $\mid$ ” represents the word “given” or “if”. This is an important result, called the Multiplication Rule, which will appear again in later sections.

#### Summary 8.3.1: Multiplication Rule

$$P(E \cap F) = P(E) \cdot P(F \mid E)$$

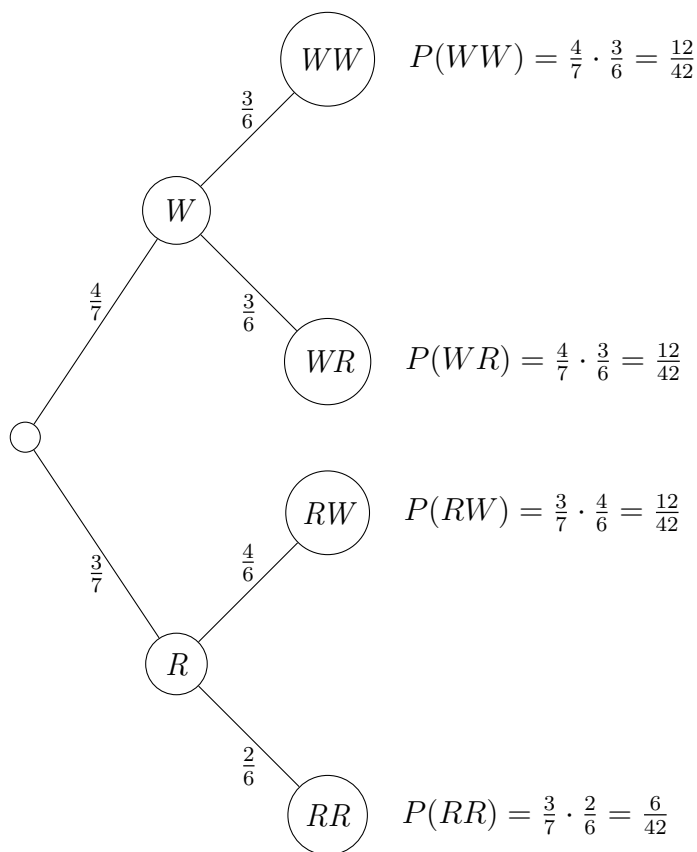
We now demonstrate the above results with a tree diagram.

**Example 8.3.3.** Suppose a jar contains 3 red and 4 white marbles. If two marbles are drawn without replacement, find the following probabilities using a tree diagram:

1. The probability that both marbles are red.
2. The probability that the first marble is red and the second white.
3. The probability that one marble is red and the other white.

**Solution 8.3.3.** Let  $R$  be the event that the marble drawn is red, and let  $W$  be the event that the marble drawn is white. We draw the following tree diagram:





1. The probability that both marbles are red is

$$P(RR) = \frac{6}{42}$$

2. The probability that the first marble is red and the second is white is

$$P(RW) = \frac{12}{42}$$

3. For the probability that one marble is red and the other is white, we observe that this can be satisfied if the first is red and the second is white, or if the first is white and the second is red. The “or” tells us we’ll be using the Addition Rule from Section 8.2.

Furthermore events  $RW$  and  $WR$  are mutually exclusive events, so we

use the form of the Addition Rule that applies to mutually exclusive events.

Therefore

$$\begin{aligned}
 &P(\text{one marble is red and the other marble is white}) \\
 &= P(RW \text{ or } WR) = P(RW) + P(WR) \\
 &= \frac{12}{42} + \frac{12}{42} \\
 &= \frac{24}{42}.
 \end{aligned}$$

### 8.3.2 Using Combinations to Find Probabilities

Although the tree diagrams give us better insight into a problem, they are not practical for problems where more than two or three things are chosen. In such cases, we use the concept of combinations that we learned in the last chapter. This method is best suited for problems where the order in which the objects are chosen is not important, and the objects are chosen without replacement.

**Example 8.3.4.** Suppose a jar contains 3 red, 2 white, and 3 blue marbles. If three marbles are drawn without replacement, find the following probabilities.

1.  $P(\text{Two red and one white})$
2.  $P(\text{One of each color})$
3.  $P(\text{None blue})$
4.  $P(\text{At least one blue})$

**Solution 8.3.4.** Let us suppose the marbles are labeled as

$$R_1, R_2, R_3, W_1, W_2, B_1, B_2, B_3.$$

1. Since we are choosing 3 marbles from a total of 8, there are  $8C3$  possible combinations. Of these  $8C3$  combinations, there are  $3C2 \cdot 2C1$  combinations consisting of 2 red and one white. Thus,

$$P(\text{Two red and one white}) = \frac{3C2 \cdot 2C1}{8C3} = \frac{6}{56} = \frac{3}{28}.$$

### 8.3. PROBABILITY USING TREE DIAGRAMS AND COMBINATIONS 243

2. For one of each color, there are  $8C3$  possible combinations. Of these 56 combinations, there are  $3C1 \cdot 2C1 \cdot 3C1 = 18$  combinations consisting of one red, one white, and one blue. we have

$$P(\text{One of each color}) = \frac{3C1 \cdot 2C1 \cdot 3C1}{8C3} = \frac{18}{56} = \frac{9}{28}$$

.

3. For none blue, there are 5 non-blue marbles, so

$$P(\text{None blue}) = \frac{5C3}{8C3} = \frac{10}{56} = \frac{5}{28}$$

.

4. By “at least one blue marble”, we mean the following: one blue marble and two non-blue marbles, or two blue marbles and one non-blue marble, or all three blue marbles. So we have to find the sum of the probabilities of all three cases.

$$P(\text{At least one blue}) = P(1 \text{ blue, 2 non-blue}) + P(2 \text{ blue, 1 non-blue}) + P(3 \text{ blue})$$

$$P(\text{At least one blue}) = \frac{3C1 \cdot 5C2}{8C3} + \frac{3C2 \cdot 5C1}{8C3} + \frac{3C3}{8C3}$$

$$P(\text{At least one blue}) = \frac{30}{56} + \frac{15}{56} + \frac{1}{56} = \frac{46}{56} = \frac{23}{28}.$$

Alternatively, for at least one blue, we can use the complement rule where  $P(E) = 1 - P(\bar{E})$  because the opposite of “at least one blue” is “none blue”. Given  $P(\text{None blue})$  from our previous work, we have

$$P(\text{At least one blue}) = 1 - P(\text{None blue}) = 1 - \frac{5}{28} = \frac{23}{28}.$$

**Example 8.3.5.** Five cards are drawn from a deck. Find the probability of obtaining two pairs, that is, two cards of one value, two of another value, and one other card.

**Solution 8.3.5.** Let us first do an easier problem—the probability of obtaining a pair of kings and queens. Since there are four kings, and four queens in the deck, the probability of obtaining two kings, two queens and one other card is

$$P(\text{A pair of kings and queens}) = \frac{4C2 \cdot 4C2 \cdot 44C1}{52C5}$$

To find the probability of obtaining two pairs, we have to consider all possible pairs. Since there are altogether 13 values, that is, aces, deuces, and so on, there are  $13C2$  different combinations of pairs.

$$P(\text{Two pairs}) = 13C2 \cdot \frac{4C2 \cdot 4C2 \cdot 44C1}{52C5} = .04754$$

**Example 8.3.6.** A cell phone store receives a shipment of 15 cell phones that contains 8 iPhones and 7 Android phones. Suppose that 6 cell phones are randomly selected from this shipment. Find the probability that a randomly selected set of 6 cell phones consists of 2 iPhones and 4 Android phones.

**Solution 8.3.6.** There are  $8C2$  ways of selecting 2 out of the 8 iPhones and  $7C4$  ways of selecting 4 out of the 7 Android phones. But altogether there are  $15C6$  ways of selecting 6 out of 15 cell phones. Therefore, we have

$$P(2\text{iPhones and } 4 \text{ Android}) = \frac{8C2 \cdot 7C4}{15C6} = \frac{(28)(35)}{5005} = \frac{980}{5005} = 0.1958$$

**Example 8.3.7.** One afternoon, a bagel store still has 53 bagels remaining: 20 plain, 15 poppyseed, and 18 sesame seed bagels. Suppose that the store owner packages up a bag of 9 bagels to bring home for tomorrow's breakfast, and selects the bagels randomly. Find the probability that the bag contains 4 plain, 3 poppyseed, and 2 sesame seed.

**Solution 8.3.7.** There are  $20C4$  ways of selecting 4 out of the 20 plain bagels,  $15C3$  ways of selecting 3 out of the 15 poppyseed bagels, and  $18C2$  ways of selecting 2 out of the 18 sesame seed bagels. But altogether there are  $53C9$  ways of selecting 9 out of the 53 bagels. Therefore, we have

$$\begin{aligned} &P(4 \text{ plain, } 3 \text{ poppyseed, and } 2 \text{ sesame seed}) \\ &= \frac{20C4 \cdot 15C3 \cdot 18C2}{53C9} = \frac{(4845)(45)(153)}{443161350} = 0.761 \end{aligned}$$

We end the section by solving a famous problem called the Birthday Problem.

**Example 8.3.8.** *If there are 25 people in a room, what is the probability that at least two people have the same birthday?*

**Solution 8.3.8.** *Let event  $E$  represent that at least two people have the same birthday.*

*We first find the probability that no two people have the same birthday.*

*We analyze as follows. Suppose there are 365 days to every year. According to the multiplication axiom, there are  $365^{25}$  possible birthdays for 25 people. Therefore, the sample space has  $365^{25}$  elements. We are interested in the probability that no two people have the same birthday. There are 365 possible choices for the first person and since the second person must have a different birthday, there are 364 choices for the second, 363 for the third, and so on. Therefore,*

$$P(\text{No two have the same birthday}) = \frac{365 \cdot 364 \cdot \dots \cdot 341}{365^{25}} = \frac{365P_{25}}{365^{25}}$$

*Since  $P(\text{at least two people have the same birthday}) = 1 - P(\text{No two have the same birthday})$ ,*

$$P(\text{at least two people have the same birthday}) = 1 - \frac{365P_{25}}{365^{25}} \approx .5687$$

This means that a teacher can supplement their meager income by betting classes of 25 students \$100 that at least two of them share a birthday. We call the amount a teacher can expect to win on average an “expected value”. In this case, the expected value is \$13.74. This is discussed further in [Section 9.3](#).

## 8.4 Conditional Probability

In this section, you will learn to:

1. recognize situations involving conditional probability
2. calculate conditional probabilities

Suppose a friend asks you the probability that it will snow today. If you are in Boston, Massachusetts in the winter, the probability of snow today might be quite substantial. If you are in Columbia, California in summer, the probability of snow today is very tiny; this probability is pretty much 0.

Let  $A$  be the event that it will snow today,  $B$  be the event that today you are in Boston in wintertime, and  $C$  be the event that today you are in Columbia in summertime.

Because the probability of snow is affected by the location and time of year, we can't just write  $P(A)$  for the probability of snow. We need to indicate the other information we know – location and time of year. We need to use conditional probability. The event we are interested in is event  $A$  for snow. The other event is called the condition, representing location and time of year in this case.

We represent conditional probability using a vertical line  $|$  that means “if”, or “given that”, or “if we know that”. The event of interest appears on the left of the  $|$ . The condition appears on the right side of the  $|$ .

The probability it will snow given that you are in Boston in the winter is represented by  $P(A|B)$ . In this case, the condition is  $B$ . The probability that it will snow given that you are in Columbia in the summer is represented by  $P(A|C)$ . In this case, the condition is  $C$ .

Now, let's examine a situation where we can calculate some probabilities.

**Example 8.4.1.** *Suppose you and a friend play a game that involves choosing a single card from a well-shuffled deck. Your friend deals you one card, face down, from the deck and offers you the following deal: If the card is a king, he will pay you \$5; otherwise, you pay him \$1. Should you play the game?*

**Solution 8.4.1.** *You reason in the following manner. Since there are four kings in the deck, the probability of obtaining a king is  $4/52$  or  $1/13$ . So, the probability of not obtaining a king is  $12/13$ . This implies that the ratio of your winning to losing is 1 to 12, while the payoff ratio is only 1 to 5. Therefore, you determine that you should not play.*

But consider the following scenario.

**Example 8.4.2.** *While your friend was dealing the card in example 8.4.1, you happened to get a glance of it and noticed that the card was a face card.*

*Should you, now, play the game?*

**Solution 8.4.2.** *Since there are 12 face cards in the deck, the total elements in the sample space are no longer 52, but just 12. This means the chance of obtaining a king is  $4/12$  or  $1/3$ . So your chance of winning is  $1/3$  and of losing  $2/3$ . This makes your winning to losing ratio 1 to 2 which fares much better with the payoff ratio of 1 to 5. This time, you determine that you should play.*

In the second part of the above example, we were finding the probability of obtaining a king knowing that a face card had shown. This is an example of conditional probability. Whenever we are finding the probability of an event  $E$  under the condition that another event  $F$  has happened, we are finding conditional probability.

The symbol  $P(E|F)$  denotes the problem of finding the probability of  $E$  given that  $F$  has occurred. We read  $P(E|F)$  as "the probability of  $E$ , given  $F$ ."

**Example 8.4.3.** *A person flips a coin three times. Find the conditional probability of getting two heads and one tail given that the first flip is a head.*

**Solution 8.4.3.** *Let event  $E$  be that the person gets two heads and one tail, and  $F$  that the first flip is a head.*

*First, we define the sample space for flipping a coin three times as follows.*

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

*Since we know the first flip is a head, our possibilities narrow down to four outcomes:  $HHH, HHT, HTH$ , and  $HTT$ .*

*Among the four,  $HHT$  and  $HTH$  represent two heads and one tail.*

*Therefore,  $P(E|F) = \frac{2}{4}$  or  $\frac{1}{2}$ .*

**Example 8.4.4.** *One six sided die is rolled once.*

- 1. Find the probability that the result is even.*
- 2. Find the probability that the result is even given that the result is greater than three.*

**Solution 8.4.4.** The sample space is  $S = \{1, 2, 3, 4, 5, 6\}$

Let event  $E$  be that the result is even and  $T$  be that the result is greater than 3.

1.  $P(E) = \frac{3}{6}$  because  $E = \{2, 4, 6\}$
2. Because  $T = \{4, 5, 6\}$ , we know that 1, 2, 3 cannot occur; only outcomes 4, 5, 6 are possible. Therefore, of the values in  $E$ , only 4, 6 are possible. Therefore,  $P(E|T) = \frac{2}{3}$

**Example 8.4.5.** A fair coin is tossed twice.

1. Find the probability that the result is two heads.
2. Find the probability that the result is two heads given that at least one head is obtained.

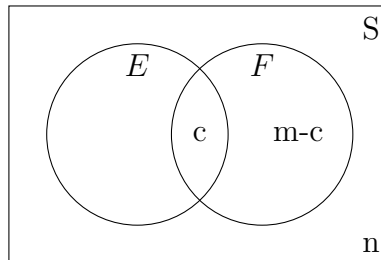
**Solution 8.4.5.** The sample space is  $S = \{HH, HT, TH, TT\}$

Let event  $E$  be that the two heads are obtained and  $F$  be at least one head is obtained

1.  $P(E) = \frac{1}{4}$  because  $E = \{HH\}$  and the sample space  $S$  has 4 outcomes.
2.  $F = \{HH, HT, TH\}$ . Since at least one head was obtained,  $TT$  did not occur. We are interested in the probability event  $E = \{HH\}$  out of the 3 outcomes in the reduced sample space  $F$ . Therefore,  $P(E|F) = \frac{1}{3}$

Let us now develop a formula for the conditional probability  $P(E|F)$ .

Suppose an experiment consists of  $n$  equally likely events. Further suppose that there are  $m$  elements in  $F$ , and  $c$  elements in  $E \cap F$ , as shown in the following Venn diagram.



If the event  $F$  has occurred, the set of all possible outcomes is no longer the entire sample space, but instead, the subset  $F$ . Therefore, we only look at the



set  $F$  and at nothing outside of  $F$ . Since  $F$  has  $m$  elements, the denominator in the calculation of  $P(E|F)$  is  $m$ . We may think that the numerator for our conditional probability is the number of elements in  $E$ . But clearly we cannot consider the elements of  $E$  that are not in  $F$ . We can only count the elements of  $E$  that are in  $F$ , that is, the elements in  $E \cap F$ . Therefore,

$$P(E|F) = \frac{c}{m}$$

Dividing both the numerator and the denominator by  $n$ , we get

$$P(E|F) = \frac{c/n}{m/n}$$

But  $c/n = P(E \cap F)$  and  $m/n = P(F)$ .

Substituting, we derive the following formula for  $P(E|F)$ :

**Summary 8.4.1: Conditional Probability Rule**

For two events  $E$  and  $F$ , the probability of “ $E$  Given  $F$ ” is

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

**Example 8.4.6.** *A single die is rolled. Use the above formula to find the conditional probability of obtaining an even number given that a number greater than three has shown.*

**Solution 8.4.6.** *Let  $E$  be the event that an even number shows, and  $F$  be the event that a number greater than three shows. We want  $P(E|F)$ .*

*$E = \{2, 4, 6\}$  and  $F = \{4, 5, 6\}$ . Which implies,  $E \cap F = \{4, 6\}$*

*Therefore,  $P(F) = \frac{3}{6}$ , and  $P(E \cap F) = \frac{2}{6}$*

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{2/6}{3/6} = \frac{2}{3}$$

**Example 8.4.7.** The following table shows the distribution of students and faculty at a community college who take public transportation and the ones who drive to school arriving before 9am. The events are abbreviated  $S$ ,  $F$ ,  $T$ , and  $D$ . Find the following probabilities:

1.  $P(D|S)$
2.  $P(F|D)$
3.  $P(S|T)$

	Student (S)	Faculty (F)	Total
Public Transportation (T)	8	13	21
Drive (D)	39	40	79
Total	47	53	

**Solution 8.4.7.** Conditional probabilities can often be found directly from a contingency table. If the condition corresponds to only one row or only one column in the table, then you can ignore the rest of the table and read the conditional probability directly from the row or column indicated by the condition.

1. The condition is event  $S$ ; we can look at only the "Student" column of the table and ignore the rest:  $P(D|S) = \frac{39}{47}$ .
2. The condition is event  $D$ ; we can look at only the "Drive" row of the table and ignore the rest:  $P(F|D) = \frac{40}{79}$ .
3. The condition is event  $T$ ; we can look at only the "Public Transportation" row of the table and ignore the rest:  $P(S|T) = \frac{8}{21}$ .

Alternatively, we can use the conditional probability formula

$$P(E|F) = \frac{P(E \cap F)}{P(F)}.$$

1.

$$P(D|M) = \frac{P(D \cap M)}{P(M)} = \frac{39/100}{47/100} = \frac{39}{47}$$

2.

$$P(F|D) = \frac{P(F \cap D)}{P(D)} = \frac{40/100}{79/100} = \frac{40}{79}$$

3.

$$P(M|T) = \frac{P(M \cap T)}{P(T)} = \frac{8/100}{21/100} = \frac{8}{21}$$

**Example 8.4.8.** Given  $P(E) = 0.5$ ,  $P(F) = 0.7$ , and  $P(E \cap F) = 0.3$ . Find the following:

1.  $P(E|F)$ 2.  $P(F|E)$ 

**Solution 8.4.8.** We use the conditional probability formula.

1.

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{0.3}{0.7} = \frac{3}{7}$$

2.

$$P(F|E) = \frac{P(E \cap F)}{P(E)} = \frac{0.3}{0.5} = \frac{3}{5}$$

**Example 8.4.9.**  $E$  and  $F$  are mutually exclusive events such that  $P(E) = 0.4$ ,  $P(F) = 0.9$ . Find  $P(E|F)$ .

**Solution 8.4.9.**  $E$  and  $F$  are mutually exclusive, so  $P(E \cap F) = 0$ . Therefore

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{0}{0.9} = 0.$$

**Example 8.4.10.** Given  $P(F|E) = 0.5$ , and  $P(E \cap F) = 0.3$ . Find  $P(E)$ .

**Solution 8.4.10.** Using the conditional probability formula  $P(F|E) = \frac{P(E \cap F)}{P(E)}$ , we get

$$\begin{aligned} P(F|E) &= \frac{P(E \cap F)}{P(E)} \\ 0.5 &= \frac{0.3}{P(E)} \\ P(E) &= \frac{0.3}{0.5} \\ P(E) &= \frac{3}{5} \end{aligned}$$

**Example 8.4.11.** *In a series of three coin flips, find the conditional probability of flipping two heads and one tail, given that there have been at least two heads.*

**Solution 8.4.11.** *Let event  $E$  be that the person flips two heads and one tail, and let  $F$  be the probability that there have been at least two heads. We want  $P(E|F)$ .*

*We list the sample space along with the events  $E$  and  $F$ .*

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

$$E = \{HHT, HTH, THH\}$$

$$F = \{HHH, HHT, HTH, THH\}$$

$$E \cap F = \{HHT, HTH, THH\}$$

*Therefore,  $P(F) = 4/8$ , and  $P(E \cap F) = 3/8$ , and*

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{3/8}{4/8} = \frac{3}{4}.$$

**Example 8.4.12.** *At a community college 65% of the students subscribe to Amazon Prime, 50% subscribe to Netflix, and 20% subscribe to both. If a student is chosen at random, find the following probabilities:*

1. *the student subscribes to Amazon Prime given that he subscribes to Netflix*
2. *the student subscribes to Netflix given that he subscribes to Amazon Prime*

**Solution 8.4.12.** *Let  $A$  be the event that the student subscribes to Amazon Prime, and  $N$  be the event that the student subscribes to Netflix. First identify the probabilities and events given in the problem.*

$$P(A) = 0.65$$

$$P(N) = 0.50$$

$$P(A \cap N) = 0.20$$

*Then use the conditional probability rule:*

1.

$$P(A|N) = \frac{P(A \cap N)}{P(N)} = \frac{0.20}{0.50} = \frac{2}{5}$$

2.

$$P(N|A) = \frac{P(A \cap N)}{P(A)} = \frac{0.20}{0.65} = \frac{4}{13}$$

## 8.5 Independent Events

In this section, you will:

1. Define independent events.
2. Identify whether two events are independent or dependent.

In the last section, we considered conditional probabilities. In some examples, the probability of an event changed when additional information was provided. This is not always the case. The additional information may or may not alter the probability of the event.

In Example 8.5.1 we revisit the discussion we started in Examples 8.4.1 and 8.4.2 at the beginning of Section 8.4 and then contrast that with Example 8.5.2.

**Example 8.5.1.** *A card is drawn from a deck. Find the following probabilities:*

1. *The card is a king.*
2. *The card is a king given that the card is a face card.*

**Solution 8.5.1.**

1. Clearly,  $P(\text{The card is a king}) = \frac{4}{52} = \frac{1}{13}$ .
2. To find  $P(\text{The card is a king} \mid \text{The card is a face card})$ , we reason as follows: There are 12 face cards in a deck of cards. There are 4 kings in a deck of cards.

$$P(\text{The card is a king} \mid \text{The card is a face card}) = \frac{4}{12} = \frac{1}{3}.$$

The reader should observe that in the above example,

$$P(\text{The card is a king} \mid \text{The card is a face card}) \neq P(\text{The card is a king})$$

In other words, the additional information, knowing that the card selected is a face card changed the probability of obtaining a king.

**Example 8.5.2.** *A card is drawn from a deck. Find the following probabilities:*

1. *The card is a king.*
2. *The card is a king given that a red card has shown.*

**Solution 8.5.2.**

1. Clearly,  $P(\text{The card is a king}) = \frac{4}{52} = \frac{1}{13}$ .
2. To find  $P(\text{The card is a king} \mid \text{A red card has shown})$ , we reason as follows: Since a red card has shown, there are only twenty-six possibilities. Of the 26 red cards, there are two kings. Therefore,

$$P(\text{The card is a king} \mid \text{A red card has shown}) = \frac{2}{26} = \frac{1}{13}.$$

The reader should observe that in the above example,

$$P(\text{The card is a king} \mid \text{A red card has shown}) = P(\text{The card is a king})$$

In other words, the additional information, a red card has shown, did not affect the probability of obtaining a king.

**Definition 8.5.1.** *Whenever the probability of an event  $E$  is not affected by the occurrence of another event  $F$ , and vice versa, we say that the two events  $E$  and  $F$  are **independent**.*

#### Summary 8.5.1: Independent Events

Two Events  $E$  and  $F$  are independent if and only if at least one of the following two conditions is true.

1.  $P(E \mid F) = P(E)$
2.  $P(F \mid E) = P(F)$

If the events are not independent, then they are dependent.

If one of these conditions is true, then both are true.

We can use the definition of independence to determine if two events are independent.

We can use that definition to develop another way to test whether two events are independent.

Recall the conditional probability formula:

$$P(E | F) = \frac{P(E \cap F)}{P(F)}$$

Multiplying both sides by  $P(F)$ , we get

$$P(E \cap F) = P(E | F)P(F)$$

Now if the two events are independent, then by definition

$$P(E | F) = P(E)$$

Substituting,  $P(E \cap F) = P(E)P(F)$

We state it formally as follows.

#### Summary 8.5.2: Test For Independence

Two events  $E$  and  $F$  are independent if and only if

$$P(E \cap F) = P(E)P(F)$$

In the Examples 8.5.3 and 8.5.4, we'll examine how to check for independence using both methods:

- Examine the probability of intersection of events to check whether  $P(E \cap F) = P(E)P(F)$
- Examine conditional probabilities to check whether  $P(E | F) = P(E)$  or  $P(F | E) = P(F)$

We need to use only one of these methods. Both methods will always give the same results.

Use the method that seems easier based on the information given in the problem.

**Example 8.5.3.** The table below shows the distribution of color-blind people by chromosome type. It should be noted that chromosome type is not exactly determinative of sex and certainly is not determinative of gender. There is also a small portion ( $\approx .2\%$ ) of the population that have  $XXY$  or  $X$  chromosome type. See [this wikipedia article](#) for some more information. Biology is really complicated because [life finds a way](#). Thank any Biologists you come across for their efforts.

	$XY$ ( $Y$ )	$XX$ ( $X$ )	Total
Color-Blind ( $C$ )	6	1	7
Not Color-Blind ( $N$ )	46	47	93
Total	52	48	100

Where  $Y$  represents  $XY$ ,  $X$  represents  $XX$ ,  $C$  represents color-blind, and  $N$  represents not color-blind. Are the events color-blind and  $XY$  independent?

**Solution 8.5.3.** According to the Test for Independence 8.5.2,  $C$  and  $Y$  are independent if and only if  $P(C \cap Y) = P(C)P(Y)$ . From the table:

$$P(C) = \frac{7}{100}, \quad P(Y) = \frac{52}{100} \quad \text{and} \quad P(C \cap Y) = \frac{6}{100}$$

So

$$P(C)P(Y) = \left(\frac{7}{100}\right) \left(\frac{52}{100}\right) = 0.0364$$

which is **not** equal to  $P(C \cap Y) = 0.06$ . Therefore, the two events are not independent. We may say they are dependent.

Alternatively, using Summary 8.5.1  $C$  and  $Y$  are independent if and only if  $P(C|Y) = P(C)$ . From the total column  $P(C) = 0.07$ . From the male column  $P(C|Y) = \frac{6}{52} = 0.1154$ . Therefore  $P(C|Y) \neq P(C)$ , indicating that the two events are not independent.

The reason for this dependence is known. The most common form of colorblindness (red-green) is associated to a misfunction of a gene on the  $X$  chromosome, so because those with  $XY$  chromosomes have only a single copy, they are much more likely to experience colorblindness.

**Example 8.5.4.** In a city with two airports, 100 flights were surveyed. 20 of those flights departed late. 45 flights in the survey departed from airport A; 9 of those flights departed late. 55 flights in the survey departed from



airport B; 11 flights departed late. Are the events "depart from airport A" and "departed late" independent?

**Solution 8.5.4.** Let  $A$  be the event that a flight departs from airport A, and  $L$  the event that a flight departs late. We have

$$\begin{aligned} P(A \cap L) &= \frac{9}{100}, \\ P(A) &= \frac{45}{100}, \\ P(L) &= \frac{20}{100}. \end{aligned}$$

According to the Test for Independence 8.5.2, in order for  $A$  and  $L$  to be independent we must have  $P(A \cap L) = P(A)P(L)$ .

Since  $P(A \cap L) = \frac{9}{100} = 0.09$  and  $P(A)P(L) = \left(\frac{45}{100}\right)\left(\frac{20}{100}\right) = 900/10000 = 0.09$ , the two events "departing from airport A" and "departing late" are independent.

Alternatively, Summary 8.5.1 states that two events are independent if  $P(E | F) = P(E)$ .

In this problem we are given that

$$P(L | A) = \frac{9}{45} = 0.2$$

and

$$P(L) = \frac{20}{100} = 0.2$$

$P(L | A) = P(L)$ , so events "departing from airport A" and "departing late" are independent.

**Example 8.5.5.** A coin is tossed three times, and the events  $E$ ,  $F$ , and  $G$  are defined as follows:

$E$ : The coin shows a head on the first toss.

$F$ : At least two heads appear.

$G$ : Heads appear in two successive tosses.

Determine whether the following events are independent:

1.  $E$  and  $F$
2.  $F$  and  $G$
3.  $E$  and  $G$

**Solution 8.5.5.** We list the sample space, the events, their intersections, and the probabilities.

Sample space  $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

$$\begin{aligned}
 E &= \{HHH, HHT, HTH, HTT\}, & P(E) &= \frac{4}{8} = \frac{1}{2} \\
 F &= \{HHH, HHT, HTH, THH\}, & P(F) &= \frac{4}{8} = \frac{1}{2} \\
 G &= \{HHT, THH\}, & P(G) &= \frac{2}{8} = \frac{1}{4} \\
 E \cap F &= \{HHH, HHT, HTH\}, & P(E \cap F) &= \frac{3}{8} \\
 F \cap G &= \{HHT, THH\}, & P(F \cap G) &= \frac{2}{8} = \frac{1}{4} \\
 E \cap G &= \{HHT\}, & P(E \cap G) &= \frac{1}{8}
 \end{aligned}$$

1.  $E$  and  $F$  will be independent if and only if  $P(E \cap F) = P(E)P(F)$

$$P(E \cap F) = \frac{3}{8} \text{ and } P(E)P(F) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Since  $\frac{3}{8} \neq \frac{1}{4}$ , we have  $P(E \cap F) \neq P(E)P(F)$ .

Events  $E$  and  $F$  are not independent.

2.  $F$  and  $G$  will be independent if and only if  $P(F \cap G) = P(F)P(G)$

$$P(F \cap G) = \frac{1}{4} \text{ and } P(F)P(G) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$$

Since  $\frac{1}{4} \neq \frac{1}{8}$ , we have  $P(F \cap G) \neq P(F)P(G)$ .

Events  $F$  and  $G$  are not independent.

3.  $E$  and  $G$  will be independent if  $P(E \cap G) = P(E)P(G)$

$$P(E \cap G) = \frac{1}{8} \text{ and } P(E)P(G) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$$

Events  $E$  and  $G$  are independent events because  $P(E \cap G) = P(E)P(G)$ .

**Example 8.5.6.** The probability that Jaime will visit his aunt in Baltimore this year is 0.30, and the probability that he will go river rafting on the Colorado river is 0.50. If the two events are independent, what is the probability that Jaime will do both?

**Solution 8.5.6.** Let  $A$  be the event that Jaime will visit his aunt this year, and  $R$  be the event that he will go river rafting. We are given  $P(A) = 0.30$  and  $P(R) = 0.50$ , and we want to find  $P(A \cap R)$ . Since we are told that the events  $A$  and  $R$  are independent,

$$P(A \cap R) = P(A)P(R) = (0.30)(0.50) = 0.15.$$

**Example 8.5.7.** Given  $P(B | A) = 0.4$ . If  $A$  and  $B$  are independent, find  $P(B)$ .

**Solution 8.5.7.** If  $A$  and  $B$  are independent, then by definition  $P(B | A) = P(B)$

Therefore,  $P(B) = 0.4$

**Example 8.5.8.** Given  $P(A) = 0.7$ ,  $P(B | A) = 0.5$ . Find  $P(A \cap B)$ .

**Solution 8.5.8.** By definition,  $P(B | A) = \frac{P(A \cap B)}{P(A)}$

Substituting, we have

$$0.5 = \frac{P(A \cap B)}{0.7}$$

Therefore,  $P(A \cap B) = 0.35$

Alternatively, starting with  $P(B | A) = \frac{P(A \cap B)}{P(A)}$

Multiplying both sides by  $P(A)$  gives

$$P(A \cap B) = P(B | A) \cdot P(A) = (0.5)(0.7) = 0.35$$

Both solutions to Example 8.5.8 are actually the same, except that the alternative approach we delayed substituting the values into the equation until after we solved the equation for  $P(A \cap B)$ . That gives the following result:

### Summary 8.5.3: Multiplication Rule

For any events  $E$  and  $F$

$$P(E \cap F) = P(E|F)P(F) \quad \text{and} \quad P(E \cap F) = P(F|E)P(E)$$

In the special case that events  $E$  and  $F$  are **independent**

$$P(E \cap F) = P(F)P(E)$$

**Example 8.5.9.** Given  $P(A) = 0.5$ ,  $P(A \cup B) = 0.7$ , if  $A$  and  $B$  are independent, find  $P(B)$ .

**Solution 8.5.9.** The addition rule states that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Since  $A$  and  $B$  are independent,  $P(A \cap B) = P(A)P(B)$

We substitute for  $P(A \cap B)$  in the addition formula and get

$$P(A \cup B) = P(A) + P(B) - P(A)P(B)$$

By letting  $P(B) = x$ , and substituting values, we get

$$0.7 = 0.5 + x - 0.5x$$

$$0.7 = 0.5 + 0.5x$$

$$0.2 = 0.5x$$

$$0.4 = x$$

Therefore,  $P(B) = 0.4$

## Chapter 9

# More Probability and Games of Chance

In this chapter, you will learn to:

1. Find the probability of a binomial experiment.
2. Find probabilities using Bayes' Formula.
3. Find the expected value or payoff in a game of chance.
4. Find probabilities using tree diagrams.

### 9.1 Binomial Probability

In this section you will learn to:

1. Recognize when to use the binomial probability distribution
2. Derive the formula for the binomial probability distribution
3. Calculate probabilities for a binomial probability experiment

In this section, we consider problems that involve a sequence of trials, where each trial has only two outcomes, a **success** or a **failure**. These trials are independent, that is, the outcome of one does not affect the outcome of any other trial. The probability of success,  $p$ , and the probability of failure,  $1 - p$ , remains the same throughout the experiment. These problems are called

**binomial probability problems.** Since these problems were researched by Swiss mathematician Jacques Bernoulli around 1700, they are also called **Bernoulli trials**.

We give the following definition:

#### Summary 9.1.1: Binomial Experiment

A binomial experiment satisfies the following four conditions:

- There are only two outcomes, a success or a failure, for each trial.
- The same experiment is repeated several times.
- The trials are independent; that is, the outcome of a particular trial does not affect the outcome of any other trial.
- The probability of success remains the same for every trial.

This probability model will give us the tools to solve many real-life problems, such as:

- If a coin is flipped 10 times, what is the probability that it will fall heads 3 times?
- If a basketball player makes 3 out of every 4 free throws, what is the probability that he will make 7 out of 10 free throws in a game?
- If a medicine cures 80% of the people who take it, what is the probability that among the ten people who take the medicine, 6 will be cured?
- If a microchip manufacturer claims that only 4% of his chips are defective, what is the probability that among the 60 chips chosen, exactly three are defective?
- If a telemarketing executive has determined that 15% of the people contacted will purchase the product, what is the probability that among the 12 people who are contacted, 2 will buy the product?

We now consider the following example to develop a formula for finding the probability of  $k$  successes in  $n$  Bernoulli trials.

**Example 9.1.1.** *A baseball player has a batting average of .300. If he bats four times in a game, find the probability that he will have*

1. 4 hits

2. 3 hits
3. 2 hits
4. 1 hit
5. no hits.

**Solution 9.1.1.** Let  $S$  denote that the player gets a hit, and  $F$  denote that he does not get a hit. This is a binomial experiment because it meets all four conditions. First, there are only two outcomes,  $S$  or  $F$ . Clearly the experiment is repeated four times. Lastly, if we assume that the player's skillfulness to get a hit does not change each time he comes to bat, the trials are independent with a probability of .3 of getting a hit during each trial.

We draw a tree diagram to show all situations. See figure 9.1.1.

Let us first find the probability of getting, for example, two hits. We will have to consider the six possibilities,  $SSFF$ ,  $FSFF$ ,  $SFSS$ ,  $SFSF$ ,  $SFSS$ ,  $FSSF$ ,  $FSFS$ , and  $FFSS$ , as shown in the above tree diagram. We list the probabilities of each below.

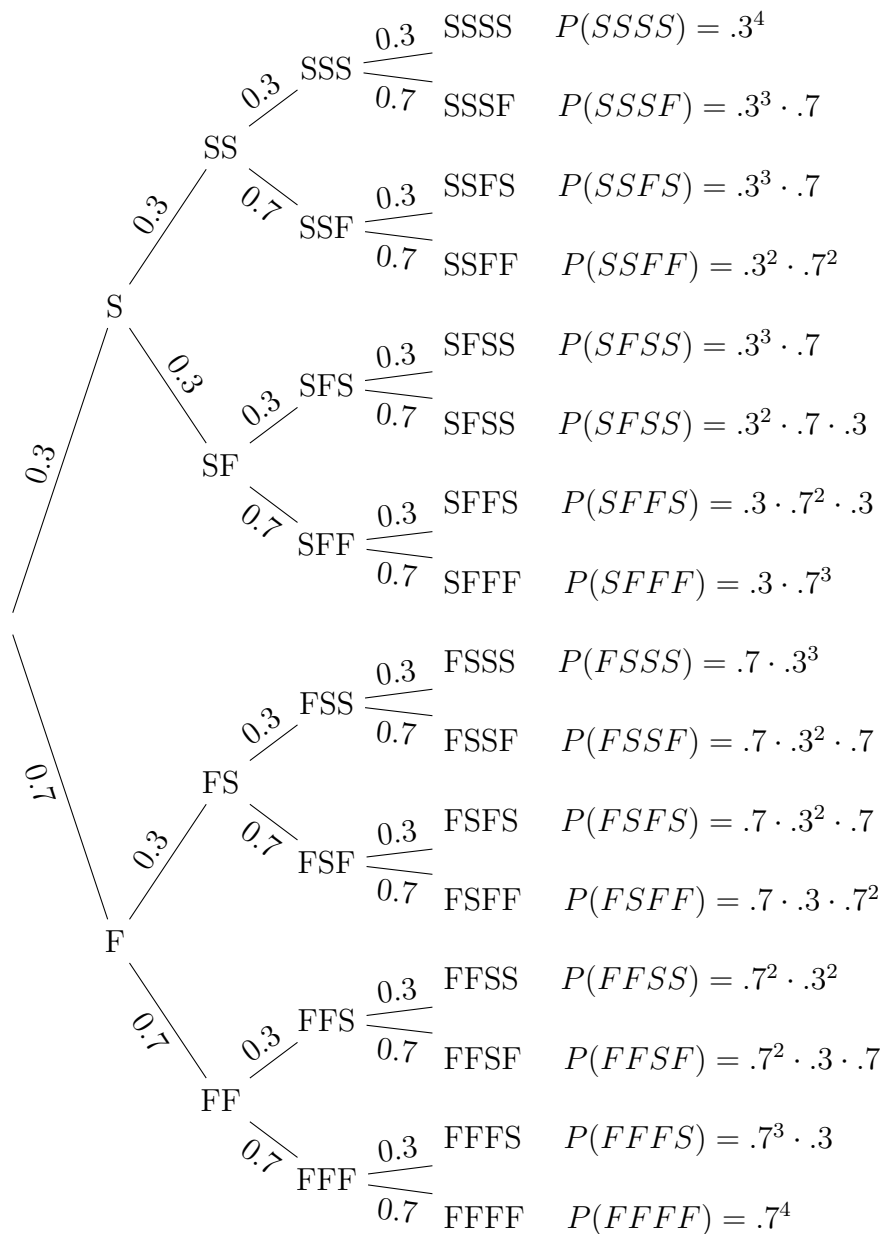
$$\begin{aligned}
 P(SSFF) &= (.3)(.3)(.7)(.7) = (.3)^2(.7)^2 \\
 P(FSFF) &= (.3)(.7)(.3)(.7) = (.3)^2(.7)^2 \\
 P(SFSF) &= (.3)(.7)(.3)(.7) = (.3)^2(.7)^2 \\
 P(SFSS) &= (.3)(.7)(.3)(.7) = (.3)^2(.7)^2 \\
 P(FSSF) &= (.7)(.3)(.3)(.7) = (.3)^2(.7)^2 \\
 P(FSFS) &= (.7)(.3)(.7)(.3) = (.3)^2(.7)^2 \\
 P(FFSS) &= (.7)(.7)(.3)(.3) = (.3)^2(.7)^2
 \end{aligned}$$

Since the probability of each of these six outcomes is  $(.3)^2(.7)^2$ , the probability of obtaining two successes is  $6(.3)^2(.7)^2$ .

The probability of getting one hit can be obtained in the same way. Since each permutation has one  $S$  and three  $F$ 's, there are four such outcomes:  $SFFF$ ,  $FSFF$ ,  $FFSF$ , and  $FFFS$ .

And since the probability of each of the four outcomes is  $(.3)(.7)^3$ , the probability of getting one hit is  $4(.3)(.7)^3$ .

Figure 9.1: Tree Diagram for 4 Bernoulli trials



The table below lists the probabilities for all cases, and shows a comparison with the binomial expansion of fourth degree. Again,  $p$  denotes the probability



of success, and  $q = (1 - p)$  the probability of failure.

<i>Outcome</i>	<i>Four Hits</i>	<i>Three Hits</i>	<i>Two Hits</i>	<i>One Hit</i>	<i>No Hits</i>
<i>Probability</i>	$.3^4$	$4(.3)^3(.7)$	$6(.3)^2(.7)^2$	$4(.3)(.7)^3$	$.7^4$

$$\begin{aligned}
 (.3 + .7)^4 &= (.3)^4 + 4(.3)^3(.7) + 6(.3)^2(.7)^2 + 4(.3)(.7)^3 + (.7)^4 \\
 (p + q)^4 &= p^4 + 4p^3q + 6p^2q^2 + 4pq^3 + q^4
 \end{aligned}$$

This gives us the following theorem:

#### Summary 9.1.2: Binomial Probability Theorem

The probability of obtaining  $k$  successes in  $n$  independent Bernoulli trials is given by

$$P(n, k; p) = {}^nC_k p^k q^{n-k}$$

where  $p$  denotes the probability of success and  $q = (1 - p)$  the probability of failure.

We use the binomial probability formula to solve the following examples.

**Example 9.1.2.** *If a coin is flipped 10 times, what is the probability that it will fall heads 3 times?*

**Solution 9.1.2.** *Let  $S$  denote the probability of obtaining a head, and  $F$  the probability of getting a tail. Clearly,  $n = 10$ ,  $k = 3$ ,  $p = \frac{1}{2}$ , and  $q = \frac{1}{2}$ . Therefore, the probability is given by*

$$P(k) = {}^{10}C_3 \cdot p^3 \cdot q^{(n-k)} = {}^{10}C_3 \cdot \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^7 = .1172$$

**Example 9.1.3.** *If a basketball player makes 3 out of every 4 free throws, what is the probability that he will make 6 out of 10 free throws in a game?*

**Solution 9.1.3.** *The probability of making a free throw is  $\frac{3}{4}$ . Therefore,  $p = \frac{3}{4}$ ,  $q = \frac{1}{4}$ ,  $n = 10$ , and  $k = 6$ . Thus, the probability is*

$$P(k) = {}^{10}C_6 \cdot p^k \cdot q^{(n-k)} = {}^{10}C_6 \cdot \left(\frac{3}{4}\right)^6 \cdot \left(\frac{1}{4}\right)^4 = .1460$$

**Example 9.1.4.** *If a medicine cures 80% of the people who take it, what is the probability that of the eight people who take the medicine, 5 will be cured?*

**Solution 9.1.4.** *Here  $p = .80$ ,  $q = .20$ ,  $n = 8$ , and  $k = 5$ .*

$$P(8, 5; .80) = (8C5)(.80)^5(.20)^3 = .1468$$

**Example 9.1.5.** *If a microchip manufacturer claims that only 4% of his chips are defective, what is the probability that among the 60 chips chosen, exactly three are defective?*

**Solution 9.1.5.** *If  $S$  denotes the probability that the chip is defective, and  $F$  the probability that the chip is not defective, then  $p = 0.04$ ,  $q = 0.96$ ,  $n = 60$ , and  $k = 3$ .*

$$P(k) = 60C3 \cdot p^3 \cdot q^{(n-k)} = 60C3 \cdot (0.04)^3 \cdot (0.96)^{57} = 0.2138.$$

**Example 9.1.6.** *A telemarketing executive has determined that 15% of people contacted will purchase the product. 12 people are contacted about this product.*

1. *Find the probability that among 12 people contacted, 2 will buy the product.*
2. *Find the probability that among 12 people contacted, at most 2 will buy the product?*

**Solution 9.1.6.** 1. *If  $S$  denotes the probability that a person will buy the product, and  $F$  the probability that the person will not buy the product, then  $p = 0.15$ ,  $q = 0.85$ ,  $n = 12$ , and  $k = 2$ .*

$$P(k) = 12C2 \cdot (0.15)^2 \cdot (0.85)^{10} = 0.2924.$$

2. *To find the probability that at most 2 people buy the product, we need to find the probabilities for  $k = 0$ ,  $k = 1$ , and  $k = 2$  and add them together.*

$$P(12, 0; .15) = 12C0 \cdot (0.15)^0 \cdot (0.85)^{12} = 0.1422$$

$$P(12, 1; .15) = 12C1 \cdot (0.15)^1 \cdot (0.85)^{11} = 0.3012$$

$$P(12, 2; .15) = 12C2 \cdot (0.15)^2 \cdot (0.85)^{10} = 0.2924$$

$$0.1422 + 0.3012 + 0.2924 = 0.7358$$

*The probability that at most 2 people buy the product is 0.7358.*

## 9.2 Bayes' Formula

In this section you will learn to:

1. Find probabilities using Bayes' formula
2. Use a probability tree to find and represent values needed when using Bayes' formula.

In this section, we will develop and use Bayes' Formula to solve an important type of probability problem. Bayes' formula is a method of calculating the conditional probability  $P(F|E)$  from  $P(E|F)$ . The ideas involved here are not new, and most of these problems can be solved using a tree diagram. However, Bayes' formula does provide us with a tool with which we can solve these problems without a tree diagram.

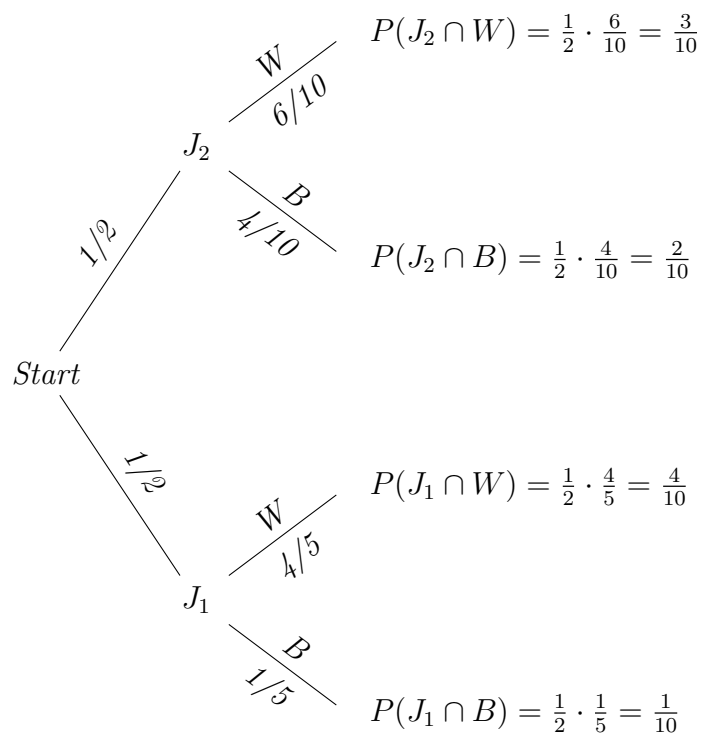
We begin with an example.

**Example 9.2.1.** *Suppose you are given two jars. Jar 1 contains one black and 4 white marbles, and Jar II contains 4 black and 6 white marbles. If a jar is selected at random and a marble is chosen,*

1. *What is the probability that the marble chosen is a black marble?*
2. *If the chosen marble is black, what is the probability that it came from Jar I?*
3. *If the chosen marble is black, what is the probability that it came from Jar II?*

**Solution 9.2.1.** *Let  $J_1$  be the event that Jar I is chosen,  $J_2$  the event that Jar 2 is chosen,  $B$  the event that a black marble is chosen and  $W$  the event that a white marble is chosen.*

*We illustrate using a tree diagram.*



1. The probability that a black marble is chosen is  $P(B) = \frac{1}{10} + \frac{2}{10} = \frac{3}{10}$ .
2. To find  $P(J_I|B)$ , we use the definition of conditional probability, and we get

$$P(J_I|B) = \frac{P(J_I \cap B)}{P(B)} = \frac{\frac{1}{10}}{\frac{3}{10}} = \frac{1}{3}$$

3. Similarly,

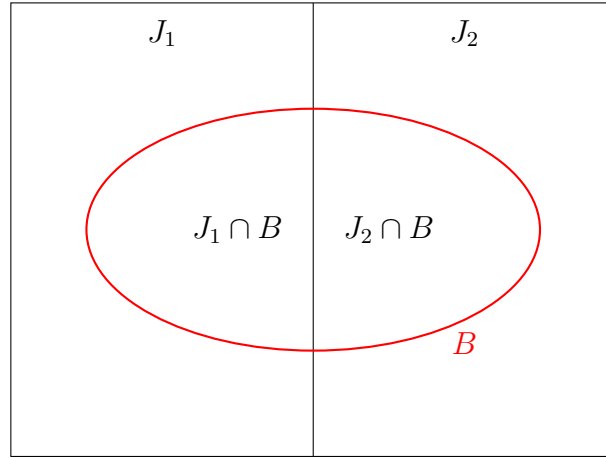
$$P(J_2|B) = \frac{P(J_2 \cap B)}{P(B)} = \frac{\frac{2}{10}}{\frac{3}{10}} = \frac{2}{3}$$

In the second and third part, the reader should note that the denominator is the sum of all probabilities of all branches of the tree that produce a black marble, while the numerator is the branch that is associated with the particular jar in question.

We will soon discover that this is a statement of Bayes' formula.

Let us first visualize the problem.

We are given a sample space  $S$  and two mutually exclusive events  $J_1$  and  $J_2$ . That is, the two events,  $J_1$  and  $J_2$ , divide the sample space into two parts such that  $J_1 \cup J_2 = S$ . Furthermore, we are given an event  $B$  that has elements in both  $J_1$  and  $J_2$ , as shown in the Venn diagram below.



From the Venn diagram, we can see that  $B = (B \cap J_1) \cup (B \cap J_2)$ . Therefore:

$$P(B) = P(B \cap J_1) + P(B \cap J_2)$$

But the multiplication rule in Chapter 8 (Summary 8.5.3) gives us

$$P(B \cap J_1) = P(J_1) \cdot P(B|J_1) \quad \text{and} \quad P(B \cap J_2) = P(J_2) \cdot P(B|J_2)$$

Substituting, we get

$$P(B) = P(J_1) \cdot P(B|J_1) + P(J_2) \cdot P(B|J_2)$$

The conditional probability rule (Summary 8.4.1) gives us

$$P(J_1|B) = \frac{P(J_1 \cap B)}{P(B)}$$

Therefore,

$$P(J_1|B) = \frac{P(J_1) \cdot P(B|J_1)}{P(B)}$$

or

$$P(J_1|B) = \frac{P(J_1) \cdot P(B|J_1)}{P(J_1) \cdot P(B|J_1) + P(J_2) \cdot P(B|J_2)}$$

The last statement is Bayes' Formula for the case where the sample space is divided into two partitions.

The following is the generalization of Bayes' formula for  $n$  partitions.

**Summary 9.2.1: Bayes' Formula**

Let  $S$  be a sample space that is divided into  $n$  partitions,  $A_1, A_2, \dots, A_n$ . If  $E$  is any event in  $S$ , then

$$P(A_i|E) = \frac{P(A_i)P(E|A_i)}{P(A_1)P(E|A_1) + P(A_2)P(E|A_2) + \dots + P(A_n)P(E|A_n)}$$

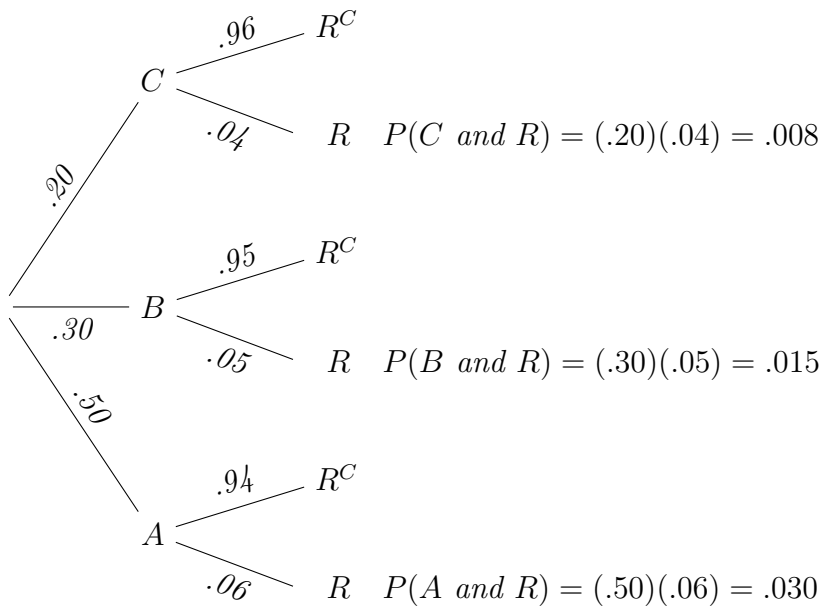
**Example 9.2.2.** *A department store buys 50% of its appliances from Manufacturer A, 30% from Manufacturer B, and 20% from Manufacturer C. It is estimated that 6% of Manufacturer A's appliances, 5% of Manufacturer B's appliances, and 4% of Manufacturer C's appliances need repair before the warranty expires. An appliance is chosen at random. If the appliance chosen needed repair before the warranty expired, what is the probability that the appliance was manufactured by Manufacturer A? Manufacturer B? Manufacturer C?*

**Solution 9.2.2.** *Let  $A$ ,  $B$  and  $C$  be the events that the appliance is manufactured by Manufacturer A, Manufacturer B, and Manufacturer C, respectively. Further, suppose that the event  $R$  denotes that the appliance needs repair before the warranty expires.*

*We need to find  $P(A|R)$ ,  $P(B|R)$  and  $P(C|R)$ .*

*We will do this problem both by using a tree diagram and by using Bayes' formula.*

*We draw a tree diagram.*



The probability  $P(A|R)$ , for example, is a fraction whose denominator is the sum of all probabilities of all branches of the tree that result in an appliance that needs repair before the warranty expires, and the numerator is the branch that is associated with Manufacturer A.  $P(B|R)$  and  $P(C|R)$  are found in the same way.

$$P(A|R) = \frac{.030}{(.030) + (.015) + (.008)} = \frac{.030}{.053} = .566$$

$$P(B|R) = \frac{.015}{.053} = .283 \quad \text{and} \quad P(C|R) = \frac{.008}{.053} = .151.$$

Alternatively, using Bayes' formula,

$$\begin{aligned} P(A|R) &= \frac{P(A)P(R|A)}{P(A)P(R|A) + P(B)P(R|B) + P(C)P(R|C)} \\ &= \frac{.030}{(.030) + (.015) + (.008)} = \frac{.030}{.053} = .566 \end{aligned}$$

$P(B|R)$  and  $P(C|R)$  can be determined in the same manner.

**Example 9.2.3.** *There are five Jacy's department stores in San Jose. The number of employees at each store and the percentage of employees that are women at each store is given in the table below. If an employee chosen at random is a woman, what is the probability that the employee works at store 3?*

Store Number	Number of Employees	Percentage Women
1	300	.40
2	150	.65
3	200	.60
4	250	.50
5	100	.70

**Solution 9.2.3.** *Let  $k = 1, 2, \dots, 5$  be the event that the employee worked at store  $k$ , and let  $W$  be the event that the employee is a woman. Since there are a total of 1000 employees at the five stores,*

$$P(1) = .30 \quad P(2) = .15 \quad P(3) = .20 \quad P(4) = .25 \quad P(5) = .10$$

*Using Bayes' formula, we calculate:*

$$\begin{aligned}
 & \frac{P(3|W)}{P(3)P(W|3)} \\
 = & \frac{P(3)P(W|3)}{P(1)P(W|1) + P(2)P(W|2) + P(3)P(W|3) + P(4)P(W|4) + P(5)P(W|5)} \\
 = & \frac{(.20)(.60)}{(.30)(.40) + (.15)(.65) + (.20)(.60) + (.25)(.50) + (.10)(.70)} \\
 = & .2254
 \end{aligned}$$

## 9.3 Expected Value

In this section you will learn to:

1. find the expected value of a discrete probability distribution
2. interpret expected value as a long-run average

**Definition 9.3.1.** *An expected gain or loss in a game of chance is called **Expected Value**.*



The concept of expected value is closely related to a weighted average. Consider the following situations:

1. Suppose you and your friend play a game that consists of rolling a die. Your friend offers you the following deal: If the die shows any number from 1 to 5, he will pay you the face value of the die in dollars; that is, if the die shows a 4, he will pay you \$4. But if the die shows a 6, you will have to pay him \$18.

Before you play the game you decide to find the expected value. You analyze as follows. Since a die will show a number from 1 to 6, with an equal probability of  $1/6$ , your chance of winning \$1 is  $1/6$ , winning \$2 is  $1/6$ , and so on up to the face value of 5. But if the die shows a 6, you will lose \$18. You write the expected value.

$$E = \$1 \left( \frac{1}{6} \right) + \$2 \left( \frac{1}{6} \right) + \$3 \left( \frac{1}{6} \right) + \$4 \left( \frac{1}{6} \right) + \$5 \left( \frac{1}{6} \right) - \$18 \left( \frac{1}{6} \right) = -\$0.50$$

This means that every time you play this game, you can expect to lose 50 cents. In other words, if you play this game 100 times, theoretically you will lose \$50. Obviously, it is not to your interest to play.

2. Suppose of the ten quizzes you took in a course, on eight quizzes you scored 80, and on two you scored 90. You wish to find the average of the ten quizzes.

$$A = \frac{(80)(8) + (90)(2)}{10} = \frac{80 \times 8 + 90 \times 2}{10} = 82$$

It should be observed that it would be incorrect to take the average of 80 and 90 because you scored 80 on eight quizzes, and 90 on only two of them. Therefore, you take a "weighted average" of 80 and 90. That is, the average of 8 parts of 80 and 2 parts of 90, which is 82.

In the first situation, to find the expected value, we multiplied each payoff by the probability of its occurrence, and then added up the amounts calculated for all possible cases. In the second example, if we consider our test score a payoff, we did the same. This leads us to the following

**Summary 9.3.1: Expected Value**

If an experiment  $X$  has the following probability distribution,

Payoff	$x_1$	$x_2$	$x_3$	$\dots$	$x_n$
Probability	$p(x_1)$	$p(x_2)$	$p(x_3)$	$\dots$	$p(x_n)$

then the **expected value** of the experiment is

$$E(X) = x_1p(x_1) + x_2p(x_2) + x_3p(x_3) + \dots + x_np(x_n)$$

**Example 9.3.1.** *In a town, 10% of the families have three children, 60% of the families have two children, 20% of the families have one child, and 10% of the families have no children. What is the expected number of children to a family?*

**Solution 9.3.1.** *We list the information in the following table:*

<i>Number of children</i>	<i>Probability</i>
3	.10
2	.60
1	.20
0	.10

*The expected value is given by  $E = x_1p(x_1) + x_2p(x_2) + x_3p(x_3) + x_4p(x_4)$ , so*

$$E = 3(.10) + 2(.60) + 1(.20) + 0(.10) = 1.7$$

*So on average, there are 1.7 children to a family.*

**Example 9.3.2.** *To sell an average house, a real estate broker spends \$1200 for advertisement expenses. If the house sells in three months, the broker makes \$8000. Otherwise, the broker loses the listing. If there is a 40% chance that the house will sell in three months, what is the expected payoff for the real estate broker?*

**Solution 9.3.2.** *The broker makes \$8000 with a probability of .40, but he loses \$1200 whether the house sells or not.*

$$\begin{aligned}
 E &= (\$8000)(.40) - (\$1200) \\
 &= \$3200 - \$1200 \\
 &= \$2000.
 \end{aligned}$$

*Alternatively, the broker makes (\$8000 - \$1200) with a probability of .40, but loses \$1200 with a probability of .60. Therefore,*

$$\begin{aligned}
 E &= (\$6800)(.40) - (\$1200)(.60) \\
 &= \$2720 - \$720 \\
 &= \$2000.
 \end{aligned}$$

**Example 9.3.3.** *In a town, the attendance at a football game depends on the weather. On a sunny day the attendance is 60,000, on a cold day the attendance is 40,000, and on a stormy day the attendance is 30,000. If for the next football season, the weatherman has predicted that 30% of the days will be sunny, 50% of the days will be cold, and 20% days will be stormy, what is the expected attendance for a single game?*

**Solution 9.3.3.** *Using the expected value formula, we get*

$$E = (60,000)(.30) + (40,000)(.50) + (30,000)(.20) = 44,000.$$

**Example 9.3.4.** *A lottery consists of choosing 6 numbers from a total of 51 numbers. The person who matches all six numbers wins \$2 million. If the lottery ticket costs \$1, what is the expected payoff?*

**Solution 9.3.4.** *Since there are  ${}_{51}C_6 = 18,009,460$  combinations of six numbers from a total of 51 numbers, the chance of choosing the winning number is 1 out of 18,009,460. So the expected payoff is:*

$$E = (\$2,000,000) \left( \frac{1}{18,009,460} \right) - \$1 = -\$0.89$$

*This means that every time a person spends \$1 to buy a ticket, he or she can expect to lose 89 cents.*

## 9.4 Probability Using Tree Diagrams

In this section you will learn to:

1. Use probability trees to organize information in probability problems
2. Use probability trees to calculate probabilities

As we have already seen, tree diagrams play an important role in solving probability problems. A tree diagram helps us not only visualize, but also list all possible outcomes in a systematic fashion. Furthermore, when we list various outcomes of an experiment and their corresponding probabilities on a tree diagram, we gain a better understanding of when probabilities are multiplied and when they are added.

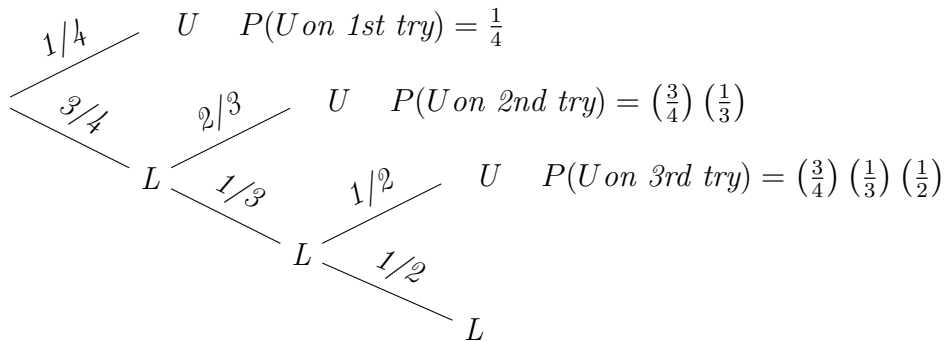
The meanings of the words "and" and "or" become clear when we learn to multiply probabilities horizontally across branches, and add probabilities vertically down the tree.

Although tree diagrams are not practical in situations where the possible outcomes become large, they are a significant tool in breaking the problem down in a schematic way. We consider some examples that may seem difficult at first, but with the help of a tree diagram, they can easily be solved.

**Example 9.4.1.** *A person has four keys and only one key fits the lock of a door. What is the probability that the locked door can be unlocked in at most three tries?*

**Solution 9.4.1.** *Let  $U$  be the event that the door has been unlocked and  $L$  be the event that the door has not been unlocked. Since there is only one correct key out of four, the probability of not unlocking the door with one try is  $\frac{3}{4}$ . The probability of not unlocking the door in three tries is  $\left(\frac{3}{4}\right)^3$ . Therefore, the probability of unlocking the door in at most three tries is the complement:*

$$P(U) = 1 - P(L) = 1 - \left(\frac{3}{4}\right)^3 = 1 - \frac{27}{64} = \frac{37}{64}$$



The probability of unlocking the door in the first try =  $\frac{1}{4}$

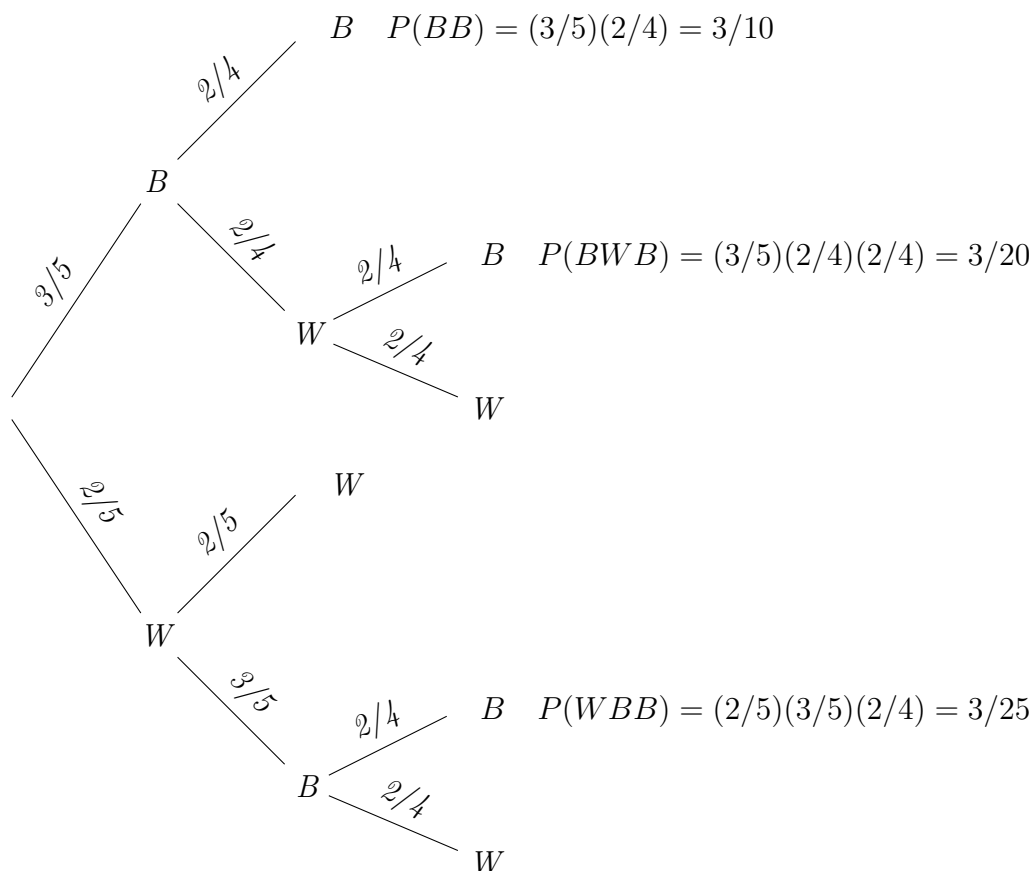
The probability of unlocking the door in the second try =  $\left(\frac{3}{4}\right) \left(\frac{1}{3}\right) = \frac{1}{4}$

The probability of unlocking the door in the third try =  $\left(\frac{3}{4}\right) \left(\frac{2}{3}\right) \left(\frac{1}{2}\right) = \frac{1}{4}$

Therefore, the probability of unlocking the door in at most three tries =  $\frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$ .

**Example 9.4.2.** A jar contains 3 black and 2 white marbles. We continue to draw marbles one at a time until two black marbles are drawn. If a white marble is drawn, the outcome is recorded and the marble is put back in the jar before drawing the next marble. What is the probability that we will get exactly two black marbles in at most three tries?

**Solution 9.4.2.** Starting with a tree diagram



The probability that we will get two black marbles in the first two tries is listed adjacent to the lowest branch, and it is  $\frac{3}{10}$ .

The probability of getting first black, second white, and third black =  $\frac{3}{20}$ .

Similarly, the probability of getting first white, second black, and third black =  $\frac{3}{25}$ .

Therefore, the probability of getting exactly two black marbles in at most three tries =  $\frac{3}{10} + \frac{3}{20} + \frac{3}{25} = \frac{57}{100}$ .

**Example 9.4.3.** A circuit consists of three resistors: resistor  $R_1$ , resistor  $R_2$ , and resistor  $R_3$ , joined in a series. If one of the resistors fails, the circuit stops working. The probabilities that resistors  $R_1$ ,  $R_2$ , or  $R_3$  will fail are 0.07, 0.10, and 0.08, respectively. Find the probability that at least one of the resistors will fail.

**Solution 9.4.3.** *The probability that at least one of the resistors fails is 1—none of the resistors fails. We don't even need to draw a tree because we can visualize the only branch of the tree that assures this outcome.*

*The probabilities that  $R_1$ ,  $R_2$ ,  $R_3$  will not fail are 0.93, 0.90, and 0.92 respectively. Therefore, the probability that none of the resistors fails is*

$$(0.93)(0.90)(0.92) = 0.77$$

*Thus, the probability that at least one of them will fail is*

$$1 - 0.77 = 0.23$$





# Chapter 10

## Markov Chains

In this chapter, you will learn to:

1. Write transition matrices for Markov Chain problems.
2. Explore some ways in which Markov Chains models are used in business, finance, public health, and other fields of application.
3. Find the long-term trend for a Regular Markov Chain.
4. Solve and interpret Absorbing Markov Chains.

### 10.1 Introduction to Markov Chains

In this section, you will learn to:

1. Write transition matrices for Markov Chain problems.
2. Use the transition matrix and the initial state vector to find the state vector that gives the distribution after a specified number of transitions.

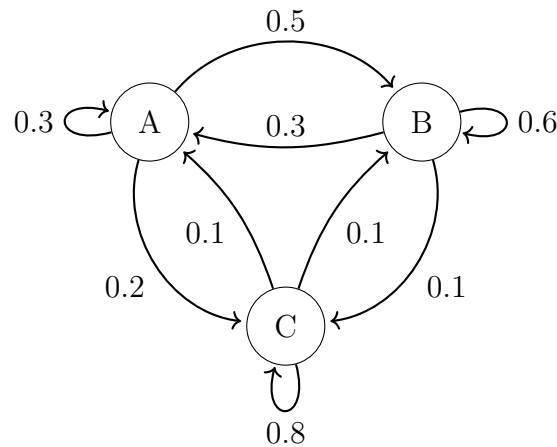
We will now study stochastic processes, experiments in which the outcomes of events depend on the previous outcomes; stochastic processes involve random outcomes that can be described by probabilities. Such a process or experiment is called a Markov Chain or Markov process. The process was first studied by a Russian mathematician named Andrei A. Markov in the early 1900s.

About 600 cities worldwide have bike share programs. Typically, a person pays a fee to join a program and can borrow a bicycle from any bike share station and then can return it to the same or another system. Each day, the distribution of bikes at the stations changes, as the bikes get returned to different stations from where they are borrowed.

For simplicity, let's consider a very simple bike share program with only 3 stations: A, B, C. Suppose that all bicycles must be returned to the station at the end of the day, so that each day there is a time, let's say midnight, that all bikes are at some station, and we can examine all the stations at this time of day, every day. We want to model the movement of bikes from midnight of a given day to midnight of the next day. We find that over a 1-day period:

- Of the bikes borrowed from station A, 30% are returned to station A, 50% end up at station B, and 20% end up at station C.
- Of the bikes borrowed from station B, 10% end up at station A, 60% have been returned to station B, and 30% end up at station C.
- Of the bikes borrowed from station C, 10% end up at station A, 10% end up at station B, and 80% are returned to station C.

We can draw an arrow diagram to show this. The arrows indicate the station where the bicycle was started, called its initial state, and the stations at which it might be located one day later, called the terminal states. The numbers on the arrows show the probability for being in each of the indicated terminal states.



Because our bike share example is simple and only has 3 stations, the arrow diagram, also called a directed graph, helps us visualize the information. But if we had an example with 10, 20, or more bike share stations, the diagram would become so complicated that it would be difficult to understand the information in the diagram.

We can use a **transition table** to organize the information (you may not need this step in the future),

	A	B	C
A	0.3	0.5	0.2
B	0.1	0.6	0.3
C	0.1	0.1	0.8

We'll use the transition table to make a **transition matrix**. Each row in the matrix represents an initial state. Each column represents a terminal state. We will assign the rows in order to stations A, B, C, and the columns in the same order to stations A, B, C. Therefore the matrix must be a square matrix, with the same number of rows as columns. The entry in row 2 column 3, for example, would show the probability that a bike that is initially at station B will be at station C one day later: that entry is 0.30, which is the probability in the diagram for the arrow that points from B to C. We use this the letter  $T$  for transition matrix.

$$T = \begin{bmatrix} 0.3 & 0.5 & 0.2 \\ 0.1 & 0.6 & 0.3 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$$

Looking at the first row that represents bikes initially at station A, we see that 30% of the bikes borrowed from station A are returned to station A, 50% end up at station B, and 20% end up at station C, after one day.

We note some properties of the transition matrix:

- $t_{ij}$  represents the entry in row  $i$  column  $j$
- $t_{ij}$  = the probability of moving from state represented by row  $i$  to the state represented by row  $j$  in a single transition
- $t_{ij}$  is a conditional probability which we can write as:  

$$t_{ij} = P(\text{next state is the state in column } j \mid \text{current state is the state in row } i)$$

- Each row adds to 1
- All entries are between 0 and 1, inclusive because they are probabilities.
- The transition matrix represents change over one transition period; in this example one transition is a fixed unit of time of one day.

**Example 10.1.1.** *A city is served by two internet service providers (ISPs), Crocks and Xminity.*

- *Due to their aggressive sales tactics, each year 40% of Crocks customers switch to Xminity; the other 60% of Crocks customers stay with Crocks.*
- *On the other hand, 30% of the Xminity customers switch to Crocks.*

*The two states in this example are Crocks and Xminity. Express the information above as a transition matrix which displays the probabilities of going from one state into another state.*

**Solution 10.1.1.** *The transition table is:*

	<i>Crocks</i>	<i>Xminity</i>
<i>Crocks</i>	.60	.40
<i>Xminity</i>	.30	.70

*so the transition matrix is given by*

$$T = \begin{bmatrix} .60 & .40 \\ .30 & .70 \end{bmatrix}$$

As previously noted, the reader should observe that a transition matrix is always a square matrix because all possible states must have both rows and columns. All entries in a transition matrix are non-negative as they represent probabilities. And, since all possible outcomes are considered in the Markov process, the sum of the row entries is always 1.

With a larger transition matrix, the ideas in Example 10.1.1 could be expanded to represent a market with more than 2 ISPs. The concepts of brand loyalty and switching between brands demonstrated in the ISPs example apply to many types of products, such as cell phone carriers, brands of regular purchases such as food or laundry detergent, brands major purchases such as cars or appliances, airlines that travelers choose when booking flights, or hotels chains that travelers choose to stay in.

The transition matrix shows the probabilities for transitions between states at two consecutive times. We need a way to represent the distribution among the states at a particular point in time. To do this we use a row matrix called a state vector. The state vector is a row matrix that has only one row; it has one column for each state. The entries show the distribution by state at a given point in time. All entries are between 0 and 1 inclusive, and the sum of the entries is 1.

For the bike share example with 3 bike share stations, the state vector is a  $1 \times 3$  matrix with 1 row and 3 columns. Suppose that when we start observing our bike share program, 30% of the bikes are at station A, 45% of the bikes are at station B, and 25% are at station C.

The initial state vector is

$$V_0 = [0.3 \quad 0.45 \quad 0.25]$$

The subscript 0 indicates that this is the initial distribution, before any transitions occur. If we want to determine the distribution after one transition, we'll need to find a new state vector that we'll call  $V_1$ . The subscript 1 indicates this is the distribution after 1 transition has occurred.

We find  $V_1$  by multiplying  $V_0$  by the transition matrix  $T$ , as follows:  $V_1 = V_0 T$ . Readers may wonder why we're multiplying the vector on the left hand side of the matrix instead of the right hand side as is usual. This notation is essentially traditional.

$$\begin{aligned} V_1 &= V_0 T \\ &= [0.30 \quad 0.45 \quad 0.25] \begin{bmatrix} 0.3 & 0.5 & 0.2 \\ 0.1 & 0.6 & 0.3 \\ 0.1 & 0.1 & 0.8 \end{bmatrix} \\ &= \begin{bmatrix} (0.30)(0.3) + (0.45)(0.1) + (0.25)(0.1) \\ (0.30)(0.5) + (0.45)(0.6) + (0.25)(0.1) \\ (0.30)(0.2) + (0.45)(0.3) + (0.25)(0.8) \end{bmatrix} \\ &= [0.16 \quad 0.445 \quad 0.395] \end{aligned}$$

After 1 day (1 transition), 16% of the bikes are at station A, 44.5% are at station B and 39.5% are at station C.

We showed the step by step work for the matrix multiplication above. In the future, we'll generally use technology to perform any necessary matrix multiplications, rather than showing the step by step work

Suppose now that we want to know the distribution of bicycles at the stations after two days. We need to find  $V_2$ , the state vector after two transitions. To find  $V_2$ , we multiply the state vector after one transition  $V_1$  by the transition matrix  $T$ .

$$V_2 = V_1 T = \begin{bmatrix} .16 & .445 & .395 \end{bmatrix} \begin{bmatrix} 0.3 & 0.5 & 0.2 \\ 0.1 & 0.6 & 0.3 \\ 0.1 & 0.1 & 0.8 \end{bmatrix} = \begin{bmatrix} .132 & .3865 & .4815 \end{bmatrix}$$

We note that  $V_1 = V_0 T$ , so  $V_2 = V_1 T = (V_0 T) T = V_0 T^2$

This gives an equivalent method to calculate the distribution of bicycles on day 2:

$$\begin{aligned} V_2 = V_0 T^2 &= \begin{bmatrix} 0.30 & 0.45 & 0.25 \end{bmatrix} \begin{bmatrix} 0.3 & 0.5 & 0.2 \\ 0.1 & 0.6 & 0.3 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}^2 \\ &= \begin{bmatrix} 0.30 & 0.45 & 0.25 \end{bmatrix} \begin{bmatrix} 0.16 & 0.47 & 0.37 \\ 0.12 & 0.44 & 0.44 \\ 0.12 & 0.19 & 0.69 \end{bmatrix} \\ &= V_2 = V_0 T^2 = \begin{bmatrix} .132 & .3865 & .4815 \end{bmatrix} \end{aligned}$$

After 2 days (2 transitions), 13.2% of the bikes are at station A, 38.65% are at station B and 48.15% are at station C.

We need to examine the following: What is the meaning of the entries in the matrix  $T^2$ ?

$$T^2 = T T = \begin{bmatrix} 0.3 & 0.5 & 0.2 \\ 0.1 & 0.6 & 0.3 \\ 0.1 & 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} 0.3 & 0.5 & 0.2 \\ 0.1 & 0.6 & 0.3 \\ 0.1 & 0.1 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.16 & 0.47 & 0.37 \\ 0.12 & 0.44 & 0.44 \\ 0.12 & 0.19 & 0.69 \end{bmatrix}$$

The entries in  $T^2$  tell us the probability of a bike being at a particular station after two transitions, given its initial station.

- Entry  $t_{13}$  in row 1 column 3 tells us that a bike that is initially borrowed from station A has a probability of 0.37 of being in station C after two transitions.
- Entry  $t_{32}$  in row 3 column 2 tells us that a bike that is initially borrowed from station C has a probability of 0.19 of being in station B after two transitions.

Similarly, if we raise transition matrix  $T$  to the  $n$ th power, the entries in  $T^n$  tells us the probability of a bike being at a particular station after  $n$  transitions, given its initial station.

And if we multiply the initial state vector  $V_0$  by  $T^n$ , the resulting row matrix  $V_n = V_0 T^n$  is the distribution of bicycles after  $n$  transitions.

**Example 10.1.2.** Refer to Example 10.1.1 for the transition matrix for market shares for subscribers to two ISPs.

1. Suppose that today  $1/4$  of customers subscribe to Crocks and  $3/4$  of customers subscribe to Xminity. After 1 year, what percent subscribe to each company?
2. Suppose instead that today 80% of customers subscribe to Crocks and 20% subscribe to Xminity. After 1 year, what percent subscribe to each company?

**Solution 10.1.2.**

1. The initial distribution given by the initial state vector is a  $1 \times 2$  matrix

$$V_0 = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \end{bmatrix} = [.25 \quad .75]$$

and the transition matrix is

$$T = \begin{bmatrix} .60 & .40 \\ .30 & .70 \end{bmatrix}$$

After 1 year, the distribution of customers is

$$V_1 = V_0 T = [.25 \quad .75] \begin{bmatrix} .60 & .40 \\ .30 & .70 \end{bmatrix} = [.375 \quad .625]$$

After 1 year, 37.5% of customers subscribe to Crocks and 62.5% to Xminity.

2. The initial distribution given by the initial state vector  $V_0 = [.8 \ .2]$ .  
Then

$$V_1 = V_0 T = [.8 \ .2] \begin{bmatrix} .60 & .40 \\ .30 & .70 \end{bmatrix} = [.54 \ .46]$$

In this case, after 1 year, 54% of customers subscribe to Crocks and 46% to Xminity.

Note that the distribution after one transition depends on the initial distribution; the distributions are different because of the different initial state vectors.

**Example 10.1.3.** Professor Symons either walks to school, or he rides his bicycle. If he walks to school one day, then the next day, he will walk or cycle with equal probability. But if he bicycles one day, then the probability that he will walk the next day is  $1/4$ . Express this information in a transition matrix.

**Solution 10.1.3.** We obtain the following transition matrix by properly placing the row and column entries. Note that if, for example, Professor Symons bicycles one day, then the probability that he will walk the next day is  $1/4$ , and therefore, the probability that he will bicycle the next day is  $3/4$ . So the transition table is:

	Walk	Bicycle
Walk	$1/2$	$1/2$
Bicycle	$1/4$	$3/4$

so the transition matrix is given by

$$T = \begin{bmatrix} .5 & .5 \\ .25 & .75 \end{bmatrix}$$

**Example 10.1.4.** In Example 10.1.3, if it is assumed that the initial day is Monday, write a matrix that gives probabilities of a transition from Monday to Wednesday.

**Solution 10.1.4.** If today is Monday, then Wednesday is two days from now, representing two transitions. We need to find the square,  $T^2$ , of the original transition matrix  $T$ , using matrix multiplication.

$$T = \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix}$$



$$\begin{aligned}
T^2 = TT &= \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix} \times \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix} \\
&= \begin{bmatrix} 1/4 + 1/8 & 1/4 + 3/8 \\ 1/8 + 3/16 & 1/8 + 9/16 \end{bmatrix} \\
&= \begin{bmatrix} 3/8 & 5/8 \\ 5/16 & 11/16 \end{bmatrix}
\end{aligned}$$

Recall that we do not obtain  $T^2$  by squaring each entry in matrix  $T$ , but obtain  $T^2$  by multiplying matrix  $T$  by itself using matrix multiplication. We represent the results in the following table.

		Wednesday	
		Walk	Bicycle
Monday	Walk	3/8	5/8
	Bicycle	5/16	11/16

We interpret the probabilities from the matrix  $T^2$  as follows:

$$P(\text{Walked Wednesday} \mid \text{Walked Monday}) = 3/8$$

$$P(\text{Bicycled Wednesday} \mid \text{Walked Monday}) = 5/8$$

$$P(\text{Walked Wednesday} \mid \text{Bicycled Monday}) = 5/16$$

$$P(\text{Bicycled Wednesday} \mid \text{Bicycled Monday}) = 11/16.$$

**Example 10.1.5.** The transition matrix for Example 10.1.3 is given below.

$$T = \begin{bmatrix} .5 & .5 \\ .25 & .75 \end{bmatrix}$$

Write the transition matrix from

1. Monday to Thursday
2. Monday to Friday

**Solution 10.1.5.**

1. In writing a transition matrix from Monday to Thursday, we are moving from one state to another in three steps. That is, we need to compute  $T^3$ .

$$T^3 = \begin{bmatrix} 11/32 & 21/32 \\ 21/64 & 43/64 \end{bmatrix}$$

2. To find the transition matrix from Monday to Friday, we are moving from one state to another in four steps. Therefore, we compute  $T^4$ .

$$T^4 = \begin{bmatrix} 43/128 & 85/128 \\ 85/256 & 171/256 \end{bmatrix}$$

It is important that the student is able to interpret the above matrix correctly. For example, the entry  $85/128$ , states that if Professor Symons walked to school on Monday, then there is  $85/128$  probability that he will bicycle to school on Friday.

There are certain Markov chains that tend to stabilize in the long run. We will examine these more deeply later in this chapter. The transition matrix we have used in the above example is just such a Markov chain. The next example deals with the long term trend or steady-state situation for that matrix.

**Example 10.1.6.** Suppose Professor Symons continues to walk and bicycle according to the transition matrix given in Example 10.1.3. In the long run, how often will he walk to school, and how often will he bicycle?

**Solution 10.1.6.** If we examine higher powers of the transition matrix  $T$ , we will find that it stabilizes.

$$T^5 = \begin{bmatrix} .333984 & .666015 \\ .333007 & .666992 \end{bmatrix}, \quad T^{10} = \begin{bmatrix} .33333397 & .66666603 \\ .33333301 & .66666698 \end{bmatrix},$$

And

$$T^{20} = \begin{bmatrix} 1/3 & 2/3 \\ 1/3 & 2/3 \end{bmatrix}, \quad T^n = \begin{bmatrix} 1/3 & 2/3 \\ 1/3 & 2/3 \end{bmatrix} \quad \text{for } n > 20.$$

*The matrix shows that in the long run, Professor Symons will walk to school 1/3 of the time and bicycle 2/3 of the time.*

*When this happens, we say that the system is in steady-state or state of equilibrium. In this situation, all row vectors are equal. If the original matrix is an  $n$  by  $n$  matrix, we get  $n$  row vectors that are all the same. We call this vector a fixed probability vector or the equilibrium vector  $E$ . In the above problem, the fixed probability vector  $E$  is  $[1/3 \ 2/3]$ . Furthermore, if the equilibrium vector  $E$  is multiplied by the original matrix  $T$ , the result is the equilibrium vector  $E$ . That is,*

$$ET = E, \quad \text{or} \quad [1/3 \ 2/3] \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix} = [1/3 \ 2/3].$$

## 10.2 Applications of Markov Chains

In this section you will learn

1. some ways in which Markov Chains models are used in business, finance, public health and other fields.

In the Section [10.1](#), we examined several applications of Markov chains. Before we proceed further in our investigations of the mathematics of Markov chains in the next sections, we take the time in this section to examine how Markov chains are used in real world applications.

In our bike share program example, we modelled the distribution of the locations of bicycles at bike share stations using a Markov chain. Markov chains have been proposed to model locations of cars distributed among multiple car rental locations for a car rental company, and locations of cars in car share programs. Markov chains models analyze package delivery schedules when packages are transported between several intermediate transport and storage locations on their way to their final destination. In these situations, Markov chains are often one part of larger mathematical models using a combination of other techniques, such as optimization to maximize profit or revenue or minimize cost using linear programming.

In our ISP example, we modelled market share in a simple example of two ISPs. Markov chains can be similarly used in market research studies for many types of products and services, to model brand loyalty and brand

transitions as we did in the ISPs model. In the field of finance, Markov chains can model investment return and risk for various types of investments.

Markov chains can model the probabilities of claims for insurance, such as life insurance and disability insurance, and for pensions and annuities. For example, for disability insurance, a much simplified model might include states of healthy, temporarily disabled, permanently disabled, recovered, and deceased; additional refinements could distinguish between disabled policyholders still in the waiting period before collecting benefits and claims actively collecting benefits.

Markov chains have been used in the fields of public health and medicine. Markov chains models of HIV and AIDS include states to model HIV transmission, progression to AIDS, and survival (living with HIV or AIDS) versus death due to AIDS. Comparing Markov chain models of HIV transmission and AIDS progression for various risk groups and ethnic groups can guide public health organizations in developing strategies for reducing risk and managing care for these various groups of people. In general, modeling transmission of various infectious diseases with Markov chains can help in determination of appropriate public health responses to monitor and slow or halt the transmission of these diseases and to determine the most efficient ways to approach treating the disease. Markov chains were heavily utilized during the COVID pandemic.

Markov chains have many health applications besides modeling spread and progression of infectious diseases. When analyzing infertility treatments, Markov chains can model the probability of successful pregnancy as a result of a sequence of infertility treatments. Another medical application is analysis of medical risk, such as the role of risk in patient condition following surgery; the Markov chain model quantifies the probabilities of patients progressing between various states of health.

Markov chains are used in ranking of websites in web searches. Markov chains model the probabilities of linking to a list of sites from other sites on that list; a link represents a transition. The Markov chain is analyzed to determine if there is a steady state distribution, or equilibrium, after many transitions. Once equilibrium is identified, the pages with the probabilities in the equilibrium distribution determine the ranking of the webpages. This is a very simplified description of how Google uses Markov chains and matrices to determine “Page rankings” as part of their search algorithms.

Of course, a real world use of such a model by Google would involve immense matrices with thousands of rows and columns. The size of such matrices requires some modifications and use of more sophisticated techniques than we study for Markov chains in this course. However the methods we study form the underlying basis for this concept. It is interesting to note that the term Page ranking does not refer to the fact that webpages are ranked, but instead is named after Google founder Larry Page, who was instrumental in developing this application of Markov chains in the area of web page search and rankings.

Markov chains are also used in quality analysis of cell phone and other communications transmissions. Transition matrices model the probabilities of certain types of signals being transmitted in sequence. Certain sequences of signals are more common and expected, having higher probabilities; on the other hand, other sequences of signals are rare and have low probabilities of occurrence. If certain sequences of signals that are unlikely to occur actually do occur, that might be an indication of errors in transmissions; Markov chains help identify the sequences that represent likely transmission errors.

## 10.3 Regular Markov Chains

In this section, you will learn to:

1. Identify Regular Markov Chains, which have an equilibrium or steady state in the long run.
2. Find the long term equilibrium for a Regular Markov Chain.

At the end of Section 10.1, we examined the transition matrix  $T$  for Professor Symons walking and biking to work Examples 10.1.3, 10.1.5, 10.1.6. As we calculated higher and higher powers of  $T$ , the matrix started to stabilize, and finally, it reached its steady-state or state of equilibrium. When that happened, all the row vectors became the same, and we called one such row vector a fixed probability vector or an equilibrium vector  $E$ . Furthermore, we discovered that  $ET = E$ .

In this section, we wish to answer the following four questions:

1. Does every Markov chain reach a state of equilibrium? Is there a way to determine if a Markov chain reaches a state of equilibrium?

2. Does the product of an equilibrium vector and its transition matrix always equal the equilibrium vector? That is, does  $ET = E$ ?
3. Can the equilibrium vector  $E$  be found without raising the matrix to higher powers?
4. Does the long-term market share distribution for a Markov chain depend on the initial market share?

### 10.3.1 Does every Markov chain reach the state of equilibrium?

No. Some Markov chains reach a state of equilibrium, but some do not. Some Markov chain transitions do not settle down to a fixed or equilibrium pattern. Therefore, we'd like to have a way to identify Markov chains that do reach a state of equilibrium.

One type of Markov chains that do reach a state of equilibrium are called regular Markov chains.

**Definition 10.3.1.** *A Markov chain is said to be a **regular Markov chain** if some power of its transition matrix  $T$  has only positive entries.*

To determine if a Markov chain is regular, we examine its transition matrix  $T$  and powers,  $T^n$ , of the transition matrix. If we find any power  $n$  for which  $T^n$  has only positive entries (no zero entries), then we know the Markov chain is regular and is guaranteed to reach a state of equilibrium in the long run.

Fortunately, we don't have to examine too many powers of the transition matrix  $T$  to determine if a Markov chain is regular; we use technology, calculators, or computers to do the calculations. There is a theorem that says that if an  $n \times n$  transition matrix represents  $n$  states, then we need only examine powers  $T^m$  up to  $m = (n - 1)^2 + 1$ .

If some power of the transition matrix  $T^m$  is going to have only positive entries, then that will occur for some power  $m \leq (n - 1)^2 + 1$ .

For example, if  $T$  is a  $3 \times 3$  transition matrix, then  $m = (n - 1)^2 + 1 = (3 - 1)^2 + 1 = 5$ .

- If we examine  $T, T^2, T^3, T^4$ , and  $T^5$ , and find that any of those matrices have only positive entries, then we know  $T$  is regular.

- If, however,  $T$ ,  $T^2$ ,  $T^3$ ,  $T^4$ , and  $T^5$  all have at least one zero entry and none of them have all positive entries, then we can stop checking. All higher powers of  $T$  will also have at least one zero entry, and  $T$  will not be regular.

**Example 10.3.1.** *Determine whether the following Markov chains are regular.*

1.  $A = \begin{bmatrix} 0 & 1 \\ .4 & .6 \end{bmatrix}$

2.  $B = \begin{bmatrix} 1 & 0 \\ .3 & .7 \end{bmatrix}$

**Solution 10.3.1.** 1. *The transition matrix  $A$  does not have all positive entries. But it is a regular Markov chain because*

$$A^2 = \begin{bmatrix} .4 & .6 \\ .24 & .76 \end{bmatrix}$$

*has only positive entries.*

2. *The matrix  $B$  is not a regular Markov chain because every power of  $B$  has an entry 0 in the first row, second column position.*

$$B = \begin{bmatrix} 1 & 0 \\ .3 & .7 \end{bmatrix}, \quad B^2 = \begin{bmatrix} 1 & 0 \\ .3 & .7 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ .3 & .7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ .51 & .49 \end{bmatrix}$$

*Since  $B$  is a  $2 \times 2$  matrix,  $m = (2 - 1)^{2+1} = 2$ . We've examined  $B$  and  $B^2$ , and discovered that neither has all positive entries. We don't need to examine any higher powers of  $B$ ;  $B$  is not a regular Markov chain.*

*In fact, we can show that all  $2 \times 2$  matrices that have a zero in the first row, second column position are not regular. Consider the following matrix  $M$ .*

$$M = \begin{bmatrix} a & 0 \\ b & c \end{bmatrix}$$

$$M^2 = \begin{bmatrix} a & 0 \\ b & c \end{bmatrix} \begin{bmatrix} a & 0 \\ b & c \end{bmatrix} = \begin{bmatrix} a^2 & 0 \\ ab + bc & b^2 + c^2 \end{bmatrix}$$

*Observe that the first row, second column entry,  $a \cdot 0 + 0 \cdot c$ , will always be zero, regardless of what power we raise the matrix to.*

### 10.3.2 Does the product of an equilibrium vector and its transition matrix always equal the equilibrium vector?

In more mathematical notation this question amounts to “does  $ET = E$ ?”. At this point, the reader may have already guessed that the answer is yes if the transition matrix is a regular Markov chain. We try to illustrate with the following example (10.1.1) from Section 10.1.

A city is served by two ISPs, Crocks and Xminity. Due to their aggressive sales tactics, each year 40% of Crocks customers switch to Xminity; the other 60% of Crocks customers stay with Crocks. On the other hand, 30% of the Xminity customers switch to Crocks and 70% of Xminity customers stay with Xminity.

The transition table is given below.

	Crocks	Xminity
Crocks	0.60	0.40
Xminity	0.30	0.70

If the initial market share for Crocks is 20% and for Xminity is 80%, we’d like to know the long term market share for each company.

Let matrix  $T$  denote the transition matrix for this Markov chain, and  $V_0$  denote the matrix that represents the initial market share. Then  $V_0$  and  $T$  are as follows:

$$V_0 = \begin{bmatrix} 0.20 & 0.80 \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} 0.60 & 0.40 \\ 0.30 & 0.70 \end{bmatrix}$$

Since each year people switch according to the transition matrix  $T$ , after one year the distribution for each company is as follows:

$$V_1 = V_0 \cdot T = \begin{bmatrix} 0.20 & 0.80 \end{bmatrix} \cdot \begin{bmatrix} 0.60 & 0.40 \\ 0.30 & 0.70 \end{bmatrix} = \begin{bmatrix} 0.36 & 0.64 \end{bmatrix}$$

After two years, the market share for each company is



$$V_2 = V_1 \cdot T = \begin{bmatrix} 0.36 & 0.64 \end{bmatrix} \cdot \begin{bmatrix} 0.60 & 0.40 \\ 0.30 & 0.70 \end{bmatrix} = \begin{bmatrix} 0.408 & 0.592 \end{bmatrix}$$

After three years the distribution is

$$V_3 = V_2 \cdot T = \begin{bmatrix} 0.408 & 0.592 \end{bmatrix} \cdot \begin{bmatrix} 0.60 & 0.40 \\ 0.30 & 0.70 \end{bmatrix} = \begin{bmatrix} 0.4224 & 0.5776 \end{bmatrix}$$

After 20 years the market share are given by  $V_{20} = V_0 T^{20} = \begin{bmatrix} 3/7 & 4/7 \end{bmatrix}$ .

After 21 years,  $V_{21} = V_{20} T^{21} = \begin{bmatrix} 3/7 & 4/7 \end{bmatrix}$ ; market shares are stable and did not change.

The market share after 20 years has stabilized to  $\begin{bmatrix} 3/7 & 4/7 \end{bmatrix}$ . This means that

$$\begin{bmatrix} 3/7 & 4/7 \end{bmatrix} \begin{bmatrix} .60 & .40 \\ .30 & .70 \end{bmatrix} = \begin{bmatrix} 3/7 & 4/7 \end{bmatrix}$$

Once the market share reaches an equilibrium state, it stays the same, that is,  $ET = E$ .

This helps us answer the next question.

### 10.3.3 Can the equilibrium vector $E$ be found without raising the transition matrix $T$ to large powers?

The answer to the second question provides us with a way to find the equilibrium vector  $E$ . The answer lies in the fact that  $ET = E$ . Since we have the matrix  $T$ , we can determine  $E$  from the statement  $ET = E$ .

Suppose  $E = \begin{bmatrix} e & 1 - e \end{bmatrix}$ , then  $ET = E$  gives us

$$\begin{bmatrix} e & 1 - e \end{bmatrix} \begin{bmatrix} .60 & .40 \\ .30 & .70 \end{bmatrix} = \begin{bmatrix} e & 1 - e \end{bmatrix}$$

$$\begin{bmatrix} (.60)e + .30(1 - e) & (.40)e + .70(1 - e) \end{bmatrix} = \begin{bmatrix} e & 1 - e \end{bmatrix}$$

$$[.30e + .30 \quad -.30e + .70] = [e \quad 1 - e]$$

Solving the system we get  $.30e + .30 = e$  which simplifies to  $e = 3/7$ .

Therefore,  $E = [3/7 \quad 4/7]$ .

As a result of our work in questions 2 and 3, we see that we have a choice of methods to find the equilibrium vector.

**Method 1: We can determine if the transition matrix  $T$  is regular. If  $T$  is regular, we know there is an equilibrium and we can use technology to find a high power of  $T$ .**

- For the question of what is a sufficiently high power of  $T$ , there is no “exact” answer.
- Select a “high power”, such as  $n = 30$ ,  $n = 50$ , or  $n = 98$ . Evaluate  $T^n$  on your calculator or with a computer. Check if  $T^{n+1} = T^n$ . If  $T^{n+1} = T^n$  and all the rows of  $T^n$  are the same, then we’ve found the equilibrium. The equilibrium vector is one row of  $T^n$ . But if you did not find equilibrium yet for a regular Markov chain, try using a higher power of  $T$ .

**Method 2: We can solve the matrix equation  $ET = E$ .**

- The disadvantage of this method is that it is a bit harder, especially if the transition matrix is larger than  $2 \times 2$ . However, it’s not as hard as it seems if  $T$  is not too large a matrix, because we can use the methods we learned in Chapter 2 to solve the system of linear equations rather than doing the algebra by hand.
- The advantage of solving  $ET = E$  as in Method 2 is that it can be used with matrices that are not regular. If a matrix is regular, it is guaranteed to have an equilibrium solution. If a matrix is not regular, then it may or may not have an equilibrium solution, and solving  $ET = E$  will allow us to prove that it has an equilibrium solution even if the matrix is not regular. (In mathematics, we say that being a regular matrix is a “sufficient” condition for having an equilibrium but is not a necessary condition.)

### 10.3.4 Does the long term market share for a Markov chain depend on the initial market share?

We will show that the final market share distribution for a Markov chain does not depend upon the initial market share. In fact, one does not even need to know the initial market share distribution to find the long term distribution. Furthermore, the final market share distribution can be found by simply raising the transition matrix to higher powers.

Consider the initial market share  $V_0 = [.20 \ .80]$  and the transition matrix from our ISP example above (originally Example 10.1.1)

$$T = \begin{bmatrix} .60 & .40 \\ .30 & .70 \end{bmatrix}$$

Recall we found  $T^n$ , for very large  $n$ , to be

$$\begin{bmatrix} 3/7 & 4/7 \\ 3/7 & 4/7 \end{bmatrix}.$$

Using our calculators, we can easily verify that for sufficiently large  $n$  (we used  $n = 30$ ),

$$V_0 T^n = [.20 \ .80] \begin{bmatrix} 3/7 & 4/7 \\ 3/7 & 4/7 \end{bmatrix} = [3/7 \ 4/7].$$

No matter what the initial market share, the product is  $[3/7 \ 4/7]$ . If instead the initial share is  $W_0 = [.10 \ .90]$ , then for sufficiently large  $n$

$$W_0 T^n = [.10 \ .90] \begin{bmatrix} 3/7 & 4/7 \\ 3/7 & 4/7 \end{bmatrix} = [3/7 \ 4/7].$$

For any distribution  $A = [a \ 1 - a]$ , for example,

$$\begin{aligned} [a \ 1 - a] \begin{bmatrix} 3/7 & 4/7 \\ 3/7 & 4/7 \end{bmatrix} &= [3/7(a) + 3/7(1 - a) \ 4/7(a) + 4/7(1 - a)] \\ &= [3/7 \ 4/7]. \end{aligned}$$

It makes sense; the entry  $3/7(a) + 3/7(1 - a)$ , for example, will always equal  $3/7$ .

**Example 10.3.2.** *Three companies, A, B, and C, compete against each other. The transition matrix  $T$  for people switching each month among them is given by the following transition table.*

	Company A	Company B	Company C
Company A	.1	.3	.6
Company B	.6	.2	.2
Company C	.1	.3	.6

*If the initial market share for the companies A, B, and C is  $[\.25 \ .35 \ .40]$ , what is the long term distribution?*

**Solution 10.3.2.** *Since the long term market share does not depend on the initial market share, we can simply find the transition matrix and raise the transition matrix to a large power and get the distribution.*

$$T = \begin{bmatrix} 0.1 & 0.3 & 0.6 \\ 0.6 & 0.2 & 0.2 \\ 0.1 & 0.3 & 0.6 \end{bmatrix}$$

$$T^{20} = \begin{bmatrix} 13/55 & 3/11 & 27/55 \\ 13/55 & 3/11 & 27/55 \\ 13/55 & 3/11 & 27/55 \end{bmatrix}$$

*In the long term, Company A has  $\frac{13}{55}$  (about 23.64%) of the market share, Company B has  $\frac{3}{11}$  (about 27.27%) of the market share, and Company C has  $\frac{27}{55}$  (about 49.09%) of the market share.*

**Summary 10.3.1: Regular Markov Chains**

A Markov chain is said to be a **Regular Markov chain** if some power of it has only positive entries.

Let  $T$  be a transition matrix for a regular Markov chain.

1. As we take higher powers of  $T$ ,  $T^n$ , as  $n$  becomes large, approaches a state of equilibrium.
2. If  $V_0$  is any distribution vector, and  $E$  an equilibrium vector, then  $V_0 T^n \rightarrow E$  as  $n \rightarrow \infty$ .
3. Each row of the equilibrium matrix  $T^\infty$  is a unique equilibrium vector  $E$  such that  $ET = E$ .
4. The equilibrium distribution vector  $E$  can be found by letting  $ET = E$ .

## 10.4 Absorbing Markov Chains

In this section, you will learn to:

1. Identify absorbing states and absorbing Markov chains.
2. Solve and interpret absorbing Markov chains.

In this section, we will study a type of Markov chain in which when a certain state is reached, it is impossible to leave that state. Such states are called absorbing states, and a Markov Chain that has at least one such state is called an Absorbing Markov chain.

Suppose you have the following transition matrix.

	$S_1$	$S_2$	$S_3$
$S_1$	.1	.3	.6
$S_2$	0	1	0
$S_3$	.3	.2	.5

The state  $S_2$  is an absorbing state because the probability of moving from state  $S_2$  to state  $S_2$  is 1, which means that if you are in state  $S_2$ , you will remain in state  $S_2$ . This is the way to identify an absorbing state: if the probability  $p_{ii}$  in row  $i$  and column  $i$  is 1, then state  $S_i$  is an absorbing state.

**Example 10.4.1.** Consider transition matrices  $A$ ,  $B$ ,  $C$  for Markov chains shown below. Which of the following Markov chains have an absorbing state?

$$A = \begin{bmatrix} .3 & .7 & 0 \\ 0 & 1 & 0 \\ .2 & .3 & .5 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} .1 & .3 & .4 & .2 \\ 0 & .2 & .1 & .7 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

**Solution 10.4.1.** The transition table for matrix  $A$  is given by

	$S_1$	$S_2$	$S_3$
$S_1$	.3	.7	0
$S_2$	0	1	0
$S_3$	.2	.3	.5

so matrix  $A$  has state  $S_2$  as an absorbing state since the state can only transition to itself and has no outbound transitions to other states. The transition probability from  $S_2$  to  $S_2$  is 1.

The transition table for matrix  $B$  is given by

	$S_1$	$S_2$	$S_3$
$S_1$	0	1	0
$S_2$	0	0	1
$S_3$	1	0	0

Matrix  $B$  does not have any absorbing states. It has cyclic transitions with no state leading to itself.

The transition table for matrix  $C$  is given by

	$S_1$	$S_2$	$S_3$	$S_4$
$S_1$	.1	.3	.4	.2
$S_2$	0	.2	.1	.7
$S_3$	0	0	1	0
$S_4$	0	0	0	1

Matrix  $C$  has two absorbing states,  $S_3$  and  $S_4$ . From state  $S_3$ , you can only remain in state  $S_3$ , and never transition to any other states. Similarly, from state  $S_4$ , you can only remain in state  $S_4$ , and never transition to any other states.

We summarize how to identify absorbing states.

**Definition 10.4.1.** A state  $S$  is an **absorbing state** in a Markov chain in the transition matrix if:

- The row for state  $S$  has one 1 and all other entries are 0 **AND**
- The entry that is 1 is on the main diagonal (row = column for that entry), indicating that we can never leave that state once it is entered.

Next, we define an absorbing Markov Chain

**Definition 10.4.2.** A Markov chain is an absorbing Markov Chain if:

- It has at least one absorbing state **AND**
- From any non-absorbing state in the Markov chain, it is possible to eventually move to some absorbing state (in one or more transitions).

**Example 10.4.2.** Consider transition matrices  $C$  and  $D$  for Markov chains shown below. Which of the following Markov chains is an absorbing Markov Chain?

$$C = \begin{bmatrix} .1 & .3 & .4 & .2 \\ 0 & .2 & .1 & .7 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ .2 & .2 & .2 & .2 & .2 \\ 0 & 0 & 0 & .3 & .7 \\ 0 & 0 & 0 & .6 & .4 \end{bmatrix}$$

**Solution 10.4.2.** Matrix  $C$  is an absorbing Markov Chain but  $D$  is not an absorbing Markov chain.

Matrix  $C$  has two absorbing states,  $S_3$  and  $S_4$ , and it is possible to get to state  $S_3$  and  $S_4$  from  $S_1$  and  $S_2$ .

Matrix  $D$  is not an absorbing Markov chain. It has two absorbing states,  $S_1$  and  $S_2$ , but it is never possible to get to either of those absorbing states from either  $S_4$  or  $S_5$ . If you are in state  $S_4$  or  $S_5$ , you always remain transitioning between states  $S_4$  or  $S_5$  and can never get absorbed into either state  $S_1$  or  $S_2$ .

In the remainder of this section, we'll examine absorbing Markov chains with two classic problems: the random drunkard's walk problem and the gambler's

ruin problem. And finally we'll conclude with an absorbing Markov model applied to a real world situation.

### 10.4.1 Drunkard's Random Walk

**Example 10.4.3.** *A man walks along a three-block portion of Main St. His house is at one end of the three-block section. A bar is at the other end of the three-block section. Each time he reaches a corner he randomly either goes forward one block or turns around and goes back one block. If he reaches home or the bar, he stays there. The four states are Home (H), Corner 1 ( $C_1$ ), Corner 2 ( $C_2$ ) and Bar (B). Write the transition matrix and identify the absorbing states. Find the probabilities of ending up in each absorbing state depending on the initial state.*

**Solution 10.4.3.** *The transition matrix is written below.*

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ .5 & 0 & .5 & 0 \\ 0 & .5 & 0 & .5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Home and the Bar are absorbing states. If the man arrives home, he does not leave. If the man arrives at the bar, he does not leave. Since it is possible to reach home or the bar from each of the other two corners on his walk, this is an absorbing Markov chain.

We can raise the transition matrix  $T$  to a high power,  $n$ . Once we find a power  $T^n$  that remains stable, it will tell us the probability of ending up in each absorbing state depending on the initial state.

$$T^{90} = T^{91} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{2}{3} & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$T^{91} = T^{90}$ ; the matrix does not change as we continue to examine higher powers. We see that in the long-run, the Markov chain must end up in an absorbing state. In the long run, the man must eventually end up at either his home or the bar.



The second row tells us that if the man is at corner  $C_1$ , then there is a  $\frac{2}{3}$  chance he will end up at home and a  $\frac{1}{3}$  chance he will end up at the bar. The third row tells us that if the man is at corner  $C_2$ , then there is a  $\frac{1}{3}$  chance he will end up at home and a  $\frac{2}{3}$  chance he will end up at the bar. Once he reaches home or the bar, he never leaves that absorbing state.

Note that while the matrix  $T^n$  for sufficiently large  $n$  has become stable and is not changing, it does not represent an equilibrium matrix. The rows are not all identical, as we found in the regular Markov chains that reached an equilibrium.

We can write a smaller “solution matrix” by retaining only rows that relate to the non-absorbing states and retaining only the columns that relate to the absorbing states. Then the solution matrix will have rows  $C_1$  and  $C_2$ , and columns  $H$  and  $B$ .

The solution matrix is

$$S = \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{bmatrix}$$

For clarity, we convert to a table

	$H$	$B$
$C_1$	$2/3$	$1/3$
$C_2$	$1/3$	$2/3$

The first row of the solution matrix shows that if the man is at corner  $C_1$ , then there is a  $2/3$  chance he will end up at home and a  $1/3$  chance he will end up at the bar.

The second row of the solution matrix shows that if the man is at corner  $C_2$ , then there is a  $1/3$  chance he will end up at home and a  $2/3$  chance he will end up at the bar.

The solution matrix does not show that eventually there is 0 probability of ending up in  $C_1$  or  $C_2$ , or that if you start in an absorbing state  $H$  or  $B$ , you stay there. The smaller solution matrix assumes that we understand these outcomes and does not include that information.

The next example is another classic example of an absorbing Markov chain. In the next example, we examine more of the mathematical details behind the concept of the solution matrix.

### 10.4.2 Gambler's Ruin Problem

**Example 10.4.4.** *A gambler has \$3,000, and she decides to gamble \$1,000 at a time at a Blackjack table in a casino in Las Vegas. She has told herself that she will continue playing until she goes broke or has \$5,000. Her probability of winning at Black Jack is .40. Write the transition matrix, identify the absorbing states, find the solution matrix, and determine the probability that the gambler will be financially ruined at a stage when she has \$2,000.*

Clearly, the states 0 and \$5,000 are the absorbing states. This makes sense because as soon as the gambler reaches 0, she is financially ruined, and the game is over. Similarly, if the gambler reaches \$5,000, she has promised herself to quit, and again, the game is over. The reader should note that  $p_{00} = 1$  and  $p_{55} = 1$ .

Since the gambler bets only \$1,000 at a time, she can raise or lower her money only by \$1,000 at a time. In other words, if she has \$2,000 now, after the next bet she can have \$3,000 with a probability of 0.40 and \$1,000 with a probability of 0.60.

	0	1K	2K	3K	4K	5K
0	1	0	0	0	0	0
1K	.60	0	.40	0	0	0
2K	0	.60	0	.40	0	0
3K	0	0	.60	0	.40	0
4K	0	0	0	.60	0	.40
5K	0	0	0	0	0	1

so

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ .60 & 0 & .40 & 0 & 0 & 0 \\ 0 & .60 & 0 & .40 & 0 & 0 \\ 0 & 0 & .60 & 0 & .40 & 0 \\ 0 & 0 & 0 & .60 & 0 & .40 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

To find the probability that the gambler will be financially ruined at a stage when she has \$2,000, we can compute the entries of the solution matrix for

the non-absorbing states, which will give us the desired probabilities.

$$T^\infty = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 195/211 & 0 & 0 & 0 & 0 & 16/211 \\ 171/211 & 0 & 0 & 0 & 0 & 40/211 \\ 135/211 & 0 & 0 & 0 & 0 & 76/211 \\ 81/211 & 0 & 0 & 0 & 0 & 130/211 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

so

	0	1K	2K	3K	4K	5K
0	1	0	0	0	0	0
1K	195/211	0	0	0	0	16/211
2K	171/211	0	0	0	0	40/211
3K	135/211	0	0	0	0	76/211
4K	81/211	0	0	0	0	130/211
5K	0	0	0	0	0	1

The solution table is often written in the following form, where the non-absorbing states are written as rows on the side, and the absorbing states as columns on the top.

	0	5K
1K	195/211	16/211
2K	171/211	40/211
3K	135/211	76/211
4K	81/211	130/211

The table lists the probabilities of getting absorbed in state 0 or state 5K starting from any of the four non-absorbing states. For example, if at any instance the gambler has \$3,000, then her probability of financial ruin is 135/211 and her probability reaching 5K is 76/211.

**Example 10.4.5.** *Solve the Gambler's Ruin Problem of Example 10.4.4 without raising the matrix to higher powers, and determine the number of bets the gambler makes before the game is over.*

**Solution 10.4.4.** *In solving absorbing states, it is often convenient to rearrange the matrix so that the rows and columns corresponding to the absorbing states are listed first. This is called the **Canonical form**. The transition matrix of Example 10.4.4 in the canonical form is listed below.*

	0	5K	1K	2K	3K	4K
0	1	0	0	0	0	0
5K	0	1	0	0	0	0
1K	.60	0	0	.40	0	0
2K	0	0	.60	0	.40	0
3K	0	0	0	.60	0	.40
4K	0	.40	0	0	0	.60

The canonical form divides the transition matrix into four sub-matrices as listed below.

	Absorbing	Non-absorbing
Absorbing	$I_k$	0
Non-absorbing	A	B

The matrix  $F = (I_n - B)^{-1}$  is called the **fundamental matrix** for the absorbing Markov chain, where  $I_n$  is an identity matrix of the same size as  $B$ . The  $i, j$ -th entry of this matrix tells us the average number of times the process is in the non-absorbing state  $j$  before absorption if it started in the non-absorbing state  $i$ .

$F = (I_n - B)^{-1}$  for our problem is listed below.

	1K	2K	3K	4K
1K	1.54	.90	.47	.19
2K	1.35	2.25	1.18	.47
3K	1.07	1.78	2.25	.90
4K	.64	1.07	1.35	1.54

The Fundamental matrix  $F$  helps us determine the average number of games played before absorption.

According to the matrix, the entry 1.78 in the row 3, column 2 position says that the gambler will play the game 1.78 times before she goes from \$3K to \$2K. The entry 2.25 in row 3, column 3 says that if the gambler now has \$3K, she will have \$3K on the average 2.25 times before the game is over.

We now address the question of how many bets will she have to make before she is absorbed, if the gambler begins with \$3K?

If we add the number of games the gambler plays in each non-absorbing state, we get the average number of games before absorption from that state. Therefore, if the gambler starts with \$3K, the average number of Black Jack games she will play before absorption is

$$1.07 + 1.78 + 2.25 + .90 = 6.0$$

That is, we expect the gambler will either have \$5,000 or nothing on the 7th bet.

Lastly, we find the solution matrix without raising the transition matrix to higher powers. The matrix  $FA$  gives us the solution matrix.

$$FA = \begin{bmatrix} 1.54 & .90 & .47 & .19 \\ 1.35 & 2.25 & 1.18 & .47 \\ 1.07 & 1.78 & 2.25 & .90 \\ .64 & 1.07 & 1.35 & 1.54 \end{bmatrix} \begin{bmatrix} .6 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & .4 \end{bmatrix} = \begin{bmatrix} .92 & .08 \\ .81 & .19 \\ .64 & .36 \\ .38 & .62 \end{bmatrix}$$

which is the same as the following matrix we obtained by raising the transition matrix to higher powers.

	0	5K
1K	195/211	16/211
2K	171/211	40/211
3K	135/211	76/211
4K	81/211	130/211

**Example 10.4.6.** At a professional school, students need to take and pass an English writing/speech class in order to get their professional degree. Students must take the class during the first quarter that they enroll. If they do not pass the class they take it again in the second semester. If they fail twice, they are not permitted to retake it again, and so would be unable to earn their degree.

Students can be in one of 4 states: passed the class (P), enrolled in the class for the first time (C), retaking the class (R) or failed twice and can not retake (F). Experience shows 70% of students taking the class for the first time pass and 80% of students taking the class for the second time pass. Write the

transition matrix and identify the absorbing states. Find the probability of being absorbed eventually in each of the absorbing states.

**Solution 10.4.5.** The absorbing states are  $P$  (pass) and  $F$  (fail repeatedly and can not retake). The transition table is shown below.

	$P$	$C$	$R$	$F$
$P$	1	0	0	0
$C$	.7	0	.3	0
$R$	.8	0	0	.2
$F$	0	0	0	1

If we raise the transition matrix  $T$  to a high power, we find that it remains stable and gives us the long-term probabilities of ending up in each of the absorbing states.

	P	C	R	F
P	1	0	0	0
C	.94	0	0	.06
R	.8	0	0	.2
F	0	0	0	1

Of students currently taking the class for the first time, 94% will eventually pass. 6% will eventually fail twice and be unable to earn their degree. Of students currently taking the class for the second time, 80% will eventually pass. 20% will eventually fail twice and be unable to earn their degree. The solution table contains the same information in abbreviated form

	P	F
C	.94	.06
R	.8	.2

Note that in this particular problem, we don't need to raise  $T$  to a "very high" power. If we find  $T^2$ , we see that it is actually equal to  $T^n$  for higher powers  $n$ .  $T^n$  becomes stable after two transitions; this makes sense in this problem because after taking the class twice, the student must have passed or is not permitted to retake it any longer. Therefore the probabilities should

not change any more after two transitions; by the end of two transitions, every student has reached an absorbing state.

#### Summary 10.4.1: Absorbing Markov Chains

1. A Markov chain is an absorbing Markov chain if it has at least one absorbing state and from any non-absorbing state in the Markov chain, it is possible to eventually move to some absorbing state (in one or more transitions). A state  $i$  is an absorbing state if once the system reaches state  $i$ , it stays in that state; that is,  $p_{ii} = 1$ .
2. If a transition matrix  $T$  for an absorbing Markov chain is raised to higher powers, it reaches an absorbing state called the solution matrix and stays there. The  $i, j$ -th entry of this matrix gives the probability of absorption in state  $j$  while starting in state  $i$ .
3. Alternately, the solution matrix can be found in the following manner.
  - (a) Express the transition matrix in the canonical form as below.

$$T = \begin{bmatrix} I_n & 0 \\ A & B \end{bmatrix}$$

where  $I_n$  is an identity matrix, and 0 is a matrix of all zeros.

- (b) The fundamental matrix  $F = (I - B)^{-1}$ . The fundamental matrix helps us find the number of games played before absorption.
  - (c)  $FA$  is the solution matrix, whose  $i, j$ -th entry gives the probability of absorption in state  $j$  while starting in state  $i$ .
4. The sum of the entries of a row of the fundamental matrix gives us the expected number of steps before absorption for the non-absorbing state associated with that row.





# Chapter 11

## Game Theory

In this chapter, you will learn to:

1. Solve strictly determined games.
2. Solve games involving mixed strategies.

### 11.1 Strictly Determined Games

In this section, you will learn to:

1. tell if a game is strictly determined.
2. find the optimal strategy for strictly determined games.
3. determine the value of a strictly determined game.

Game theory is one of the newest branches of mathematics. It first came to light when a brilliant mathematician named Dr. John von Neumann co-authored a book titled *Theory of Games and Economic Behavior* with Dr. Morgenstern. Since then, it has played an important role in decision making in business, economics, social sciences, and other fields.

In this chapter, we will study games that involve only two players. In these games, since a win for one person is a loss for the other, we refer to them as two-person zero-sum games. Although the games we will study here are

fairly simple, they will provide us with an understanding of how games work and how they are applied in practical situations. We begin with an example.

**Example 11.1.1.** *Robert and Carol decide to play a game using a dime and a quarter. Each chooses one of the two coins, puts it in their hand and closes their fist. At a given signal, they simultaneously open their fists. If the sum of the coins is less than 35 cents, Robert gets both coins, otherwise, Carol gets both coins. Write the matrix for the game, determine the optimal strategies for each player, and find the expected payoff for Robert.*

**Solution 11.1.1.** *Suppose Robert is the row player, that is, he plays the rows, and Carol is a column player. The following matrix depicts all four cases and their corresponding payoffs for Robert, with a negative value indicating a loss for Robert and a win for Carol.*

		Carol	
		Dime	Quarter
Robert	Dime	10	−10
	Quarter	−25	25

*The best strategy for Robert is to always show a quarter because this way the worst he can do is lose ten cents. And, the best strategy for Carol is to always show a quarter because that way the worst she can do is to lose ten cents. If they both follow their optimal strategies, Robert will lose ten cents each time. Therefore, the value of the game is negative ten cents for Robert.*

In the above example, since there is only one fixed optimal strategy for each player, regardless of their opponent's strategy, we say the game possesses a **pure strategy** and is **strictly determined**.

Next, we formulate a method to find the optimal strategy for each player and the value of the game. The method involves considering the worst scenario for each player.

To consider the worst situation, the row player considers the minimum value in each row, and the column player considers the maximum value in each column. Note that the maximum value really represents a minimum value for the column player because the game matrix depicts the payoffs for the row player. We list the method below.

**Summary 11.1.1: Finding the Optimal Strategy and the Value for Strictly Determined Games**

1. Put an asterisk (\*) next to the minimum entry in each row.
2. Put a box around the maximum entry in each column.
3. The entry that has both an asterisk and a box represents the value of the game and is called a saddle point.
4. The row that is associated with the saddle point represents the best strategy for the row player, and the column that is associated with the saddle point represents the best strategy for the column player.
5. A game matrix can have more than one saddle point, but all saddle points have the same value.
6. If no saddle point exists, the game is not strictly determined. Non-strictly determined games are the subject of the next section.

**Example 11.1.2.** Find the saddle points and optimal strategies for the following game.

	Column Player	
Row Player	-25	10
	10	50

**Solution 11.1.2.** We find the saddle point by placing an asterisk next to the minimum entry in each row, and by putting a box around the maximum entry in each column as shown below.

	Column Player	
Row Player	-25*	10
	<span style="border: 1px solid black;">10*</span>	<span style="border: 1px solid black;">50</span>

Since the second row, first column entry, which happens to be 10, has both an asterisk and a box, it is a saddle point. This implies that the value of the game is 10, and the optimal strategy for the row player is to always play row 2, and the optimal strategy for the column player is to always play column 1. If both players play their optimal strategies, the row player will win 10 units each time.

The row player's strategy is written as  $\begin{bmatrix} 0 & 1 \end{bmatrix}$  indicating that he will play row 1 with a probability of 0 and row 2 with a probability of 1.

Similarly, the column player's strategy is written as  $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$  implying that he will play column 1 with a probability of 1, and column 2 with a probability of 0.

## 11.2 Non-Strictly Determined Games

In this section, you will learn to:

1. find the optimal strategy for non-strictly determined games.
2. find the value of non-strictly determined games.

In this section, we study games that have no saddle points. This means that these games do not possess a pure strategy. We call these games non-strictly determined games. If the game is played only once, it will make no difference what move is made. However, if the game is played repeatedly, a mixed strategy consisting of alternating random moves can be worked out.

We consider the following example.

**Example 11.2.1.** Suppose Robert and Carol decide to play a game using a dime and a quarter. At a given signal, they simultaneously show one of the two coins. If the coins match, Robert gets both coins, but if they don't match, Carol gets both coins. Determine whether the game is strictly determined.

**Solution 11.2.1.** We write the payoff table for Robert as follows:

		Carol	
		Dime	Quarter
Robert	Dime	10	-10
	Quarter	-25	25

To determine whether the game is strictly determined, we look for a saddle point. Again, we place an asterisk next to the minimum entry in each row, and a box around the maximum value in each column. We get

		Carol	
		Dime	Quarter
Robert	Dime	<span style="border: 1px solid black;">10</span>	-10*
	Quarter	-25*	<span style="border: 1px solid black;">25</span>

Since there is no entry that has both an asterisk and a box, the game does not have a saddle point, and thus it is non-strictly determined.

We wish to devise a strategy for Robert. If Robert consistently shows a dime, for example, Carol will see the pattern and will start showing a quarter, and Robert will lose. Conversely, if Carol repeatedly shows a quarter, Robert will start showing a quarter, thus resulting in Carol's loss. So a good strategy is to throw your opponent off by showing a dime some of the times and showing a quarter other times. Before we develop an optimal strategy for each player, we will consider an arbitrary strategy for each and determine the corresponding payoffs.

**Example 11.2.2.** Suppose in Example 11.2.1, Robert decides to show a dime with .20 probability and a quarter with .80 probability, and Carol decides to show a dime with .70 probability and a quarter with .30 probability. What is the expected payoff for Robert?

**Solution 11.2.2.** Let  $R$  denote Robert's strategy and  $C$  denote Carol's strategy. Since Robert is a row player and Carol is a column player, their strategies are written as follows:

$$R = \begin{bmatrix} .20 & .80 \end{bmatrix}, \quad C = \begin{bmatrix} .70 \\ .30 \end{bmatrix}$$

To find the expected payoff, we use the following reasoning. Since Robert chooses to play row 1 with .20 probability and Carol chooses to play column 1 with .70 probability, the move row 1, column 1 will be chosen with  $(.20)(.70) = .14$  probability. The fact that this move has a payoff of 10 cents for Robert, Robert's expected payoff for this move is  $(.14)(10) = .014$  cents. Similarly, we compute Robert's expected payoffs for the other cases. The table below lists expected payoffs for all four cases.

Move	Probability	Payoff	Expected Payoff
Row 1, Column 1	$(.20)(.70) = .14$	10 cents	1.4 cents
Row 1, Column 2	$(.20)(.30) = .06$	-10 cents	-.6 cents
Row 2, Column 1	$(.80)(.70) = .56$	-25 cents	-14 cents
Row 2, Column 2	$(.80)(.30) = .24$	25 cents	6.0 cents
Totals	1		-7.2 cents

The above table shows that if Robert plays the game with the strategy  $R =$

$[.20 \ .80]$  and Carol plays with the strategy  $C = \begin{bmatrix} .70 \\ .30 \end{bmatrix}$ , Robert can expect to lose 7.2 cents for every game.

Alternatively, if we call the game matrix  $G$ , then the expected payoff for the row player can be determined by multiplying matrices  $R$ ,  $G$  and  $C$ . Thus, the expected payoff  $P$  for Robert is as follows:

$$P = RGC = [.20 \ .80] \begin{bmatrix} 10 & -10 \\ -25 & 25 \end{bmatrix} \begin{bmatrix} .70 \\ .30 \end{bmatrix} = -7.2 \text{ cents.}$$

which is the same as the one obtained from the table.

**Example 11.2.3.** For the following game matrix  $G$ , determine the optimal strategy for both the row player and the column player, and find the value of the game.

$$G = \begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix}$$

**Solution 11.2.3.** Let us suppose that the row player uses the strategy  $R = [r \ 1 - r]$ . Now if the column player plays column 1, the expected payoff  $P$  for the row player is:

$$P(r) = 1(r) + (-3)(1 - r) = 4r - 3.$$

This can also be computed as follows:

$$P(r) = [r \ 1 - r] \begin{bmatrix} 1 \\ -3 \end{bmatrix} = 4r - 3.$$

If the row player plays the strategy  $[r \ 1 - r]$  and the column player plays column 2, the expected payoff  $P$  for the row player is

$$P(r) = [r \ 1 - r] \begin{bmatrix} -2 \\ 4 \end{bmatrix} = -6r + 4.$$

We have two equations:  $P(r) = 4r - 3$  and  $P(r) = -6r + 4$

The row player is trying to improve upon his worst scenario, and that only happens when the two lines intersect. Any point other than the point of

intersection will not result in optimal strategy as one of the expectations will fall short.

Solving for  $r$  algebraically, we get

$$4r - 3 = -6r + 4$$

$$r = \frac{7}{10}.$$

Therefore, the optimal strategy for the row player is  $[\text{.7} \text{ .3}]$ .

Alternatively, we can find the optimal strategy for the row player by, first, multiplying the row matrix with the game matrix as shown below.

$$[r \ 1-r] \begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix} = [4r-3 \ -6r+4]$$

And then by equating the two entries in the product matrix. Again, we get  $r = \text{.7}$ , which gives us the optimal strategy  $[\text{.7} \text{ .3}]$ .

We use the same technique to find the optimal strategy for the column player.

Suppose the column player's optimal strategy is represented by  $[c \ 1-c]$ . We first multiply the game matrix by the column matrix as shown below.

$$\begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} c \\ 1-c \end{bmatrix} = \begin{bmatrix} 3c-2 \\ -7c+4 \end{bmatrix}$$

And then equate the entries in the product matrix. We get

$$3c - 2 = -7c + 4$$

$$c = \text{.6}$$

Therefore, the column player's optimal strategy is  $[\text{.6} \text{ .4}]$ .

To find the expected value,  $V$ , of the game, we find the product of the matrices  $R$ ,  $G$ , and  $C$ .

$$V = [\text{.7} \text{ .3}] \begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} \text{.6} \\ \text{.4} \end{bmatrix} = V = -0.2$$

That is, if both players play their optimal strategies, the row player can expect to lose .2 units for every game.

**Example 11.2.4.** For the game in Example 11.2.1, determine the optimal strategy for both Robert and Carol, and find the value of the game. We rewrite the game matrix.

$$G = \begin{bmatrix} 10 & -10 \\ -25 & 25 \end{bmatrix}$$

**Solution 11.2.4.** Let  $R = [r \quad 1 - r]$  be Robert's strategy, and  $C = \begin{bmatrix} c \\ 1 - c \end{bmatrix}$  be Carol's strategy.

To find the optimal strategy for Robert, we first find the product  $RG$  as below.

$$[r \quad 1 - r] \begin{bmatrix} 10 & -10 \\ -25 & 25 \end{bmatrix} = [35r - 25 \quad -35r + 25]$$

By setting the entries equal, we get

$$35r - 25 = -35r + 25$$

resulting in  $r = \frac{5}{7}$ . Therefore, the optimal strategy for Robert is

$$R = [5/7 \quad 2/7].$$

To find the optimal strategy for Carol, we first find the following product.

$$\begin{bmatrix} 10 & -10 \\ -25 & 25 \end{bmatrix} \begin{bmatrix} c \\ 1 - c \end{bmatrix} = \begin{bmatrix} 20c - 10 \\ -50c + 25 \end{bmatrix}$$

We now set the entries equal to each other, and we get

$$20c - 10 = -50c + 25$$

resulting in  $c = \frac{1}{2}$ .

Therefore, the optimal strategy for Carol is

$$C = [1/2 \quad 1/2].$$

To find the expected value,  $V$ , of the game, we find the product  $RGC$ .

$$V = [5/7 \quad 2/7] \begin{bmatrix} 10 & -10 \\ -25 & 25 \end{bmatrix} \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = [0]$$

If both players play their optimal strategy, the value of the game is zero. In such case, the game is called **fair**.



## 11.3 Reduction by Dominance

In this section, you will learn to:

1. reduce by dominance.

Sometimes an  $m \times n$  game matrix can be reduced to a  $2 \times 2$  matrix by deleting certain rows and columns. A row can be deleted if there exists another row that will produce a payoff of an equal or better value. Similarly, a column can be deleted if there is another column that will produce a payoff of an equal or better value for the column player. The row or column that produces a better payoff for its corresponding player is said to dominate the row or column with the lesser payoff.

**Example 11.3.1.** *For the following game, determine the optimal strategy for both the row player and the column player, and find the value of the game.*

$$G = \begin{bmatrix} -2 & 6 & 4 \\ -1 & -2 & -3 \\ 1 & 2 & -2 \end{bmatrix}$$

**Solution 11.3.1.** *We first look for a saddle point and determine that none exist. Next, we try to reduce the matrix to a  $2 \times 2$  matrix by eliminating the dominated row.*

*Since every entry in row 3 is larger than the corresponding entry in row 2, row 3 dominates row 2. Therefore, a rational row player will never play row 2, and we eliminate row 2. We get*

$$\begin{bmatrix} -2 & 6 & 4 \\ 1 & 2 & -2 \end{bmatrix}$$

*Now we try to eliminate a column. Remember that the game matrix represents the payoffs for the row player and not the column player; therefore, the larger the number in the column, the smaller the payoff for the column player.*

*The column player will never play column 2, because it is dominated by column 3. Therefore, we eliminate column 2 and get the modified matrix,  $M$ ,*

$$M = \begin{bmatrix} -2 & 4 \\ 1 & -2 \end{bmatrix}$$

To find the optimal strategy for both the row player and the column player, we use the method learned in Section 11.2.

Let the row player's strategy be  $R = [r \ 1 - r]$ , and the column player's strategy be

$$C = \begin{bmatrix} c \\ 1 - c \end{bmatrix}$$

To find the optimal strategy for the row player, we, first, find the product  $RM$  as below.

$$[r \ 1 - r] \begin{bmatrix} 1 & -2 \\ 1 & 2 \end{bmatrix} = [-3r + 1 \ 6r - 2]$$

By setting the entries equal, we get

$$-3r + 1 = 6r - 2$$

or

$$r = \frac{1}{3}.$$

Therefore, the optimal strategy for the row player is  $[1/3 \ 2/3]$ , but relative to the original game matrix it is  $[1/3 \ 0 \ 2/3]$ .

To find the optimal strategy for the column player, we, first, find the following product.

$$\begin{bmatrix} -2 & 4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} c \\ 1 - c \end{bmatrix} = \begin{bmatrix} -6c + 4 \\ 3c - 2 \end{bmatrix}$$

We set the entries in the product matrix equal to each other, and we get,

$$-6c + 4 = 3c - 2$$

or

$$c = 2/3$$

Therefore, the optimal strategy for the column player is  $\begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$ , but relative to the original game matrix, the strategy for the column player is  $\begin{bmatrix} 2/3 \\ 0 \\ 1/3 \end{bmatrix}$ .

To find the expected value,  $V$ , of the game, we have two choices: either to find the product of matrices  $R$ ,  $M$  and  $C$ , or multiply the optimal strategies relative to the original matrix to the original matrix. We choose the first, and get

$$V = \begin{bmatrix} 1/3 & 2/3 \end{bmatrix} \begin{bmatrix} -2 & 4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$$

$$V = 0$$

Therefore, if both players play their optimal strategy, the value of the game is zero. In other words, the game is fair.

We summarize as follows:

#### Summary 11.3.1: Reduction by Dominance

1. Sometimes an  $m \times n$  game matrix can be reduced to a  $2 \times 2$  matrix by deleting *dominated* rows and columns.
2. A row is called a *dominated row* if there exists another row that will produce a payoff of an equal or better value. That happens when there exists a row whose every entry is larger than the corresponding entry of the dominated row.
3. A column is called a *dominated column* if there exists another column that will produce a payoff of an equal or better value. This happens when there exists a column whose every entry is smaller than the corresponding entry of the dominated row.

