CHAPTER 4

Probability

Learning Objectives

After mastering the material in this chapter, you will be able to:

- **LO4-1** Define a probability and a sample space.
- **LO4-2** List the outcomes in a sample space and use the list to compute probabilities.
- **LO4-3** Use elementary probability rules to compute probabilities.
- **LO4-4** Compute conditional probabilities and assess independence.
- **LO4-5** Use Bayes' Theorem to update prior probabilities to posterior probabilities (Optional).
- **LO4-6** Use elementary counting rules to compute probabilities (Optional).

Chapter Outline

In Chapter 3 we explained how to use sample statistics as point estimates of population parameters. Starting in Chapter 7, we will focus on using sample statistics to make more sophisticated **statistical inferences** about population parameters. We will see that these statistical inferences are generalizations—based on calculating **probabilities**—about population parameters. In this chapter and in Chapters 5 and 6 we present the fundamental concepts about probability that are needed to understand how we make such statistical inferences. We begin our discussions in this chapter by considering rules for calculating probabilities.

In order to illustrate some of the concepts in this chapter, we will introduce a new case.

C

The Crystal Cable Case: A cable company uses probability to assess the market penetration of its television and Internet services.

4.1 Probability and Sample Spaces

LO4-1 Define a probability and a sample space.

An introduction to probability and sample spaces We use the concept of **probability** to deal with uncertainty. Intuitively, the probability of an event is a number that measures the chance, or likelihood, that the event will occur. For instance, the probability that your favorite football team will win its next game measures the likelihood of a victory. The probability of an event is always a number between 0 and 1. The closer an event's probability is to 1, the higher is the likelihood that the event will occur; the closer the event's probability is to 0, the smaller is the likelihood that the event will occur. For example, if you believe that the probability that your favorite football team will win its next game is .95, then you are almost sure that your team will win. However, if you believe that the probability of victory is only .10, then you have very little confidence that your team will win.

When performing statistical studies, we sometimes collect data by **performing a controlled experiment.** For instance, we might purposely vary the operating conditions of a manufacturing process in order to study the effects of these changes on the process output. Alternatively, we sometimes obtain data by **observing uncontrolled events.** For example, we might observe the closing price of a share of General Motors' stock every day for 30 trading days. In order to simplify our terminology, we will use the word *experiment* to refer to either method of data collection. We now formally define an experiment and the *sample space* of an experiment.

An **experiment** is any process of observation that has an uncertain outcome. The **sample space** of an experiment is the set of all possible outcomes for the experiment. The possible outcomes are sometimes called **experimental outcomes** or **sample space outcomes.**

When specifying the sample space of an experiment, we must define the sample space outcomes so that on any single repetition of the experiment, one and only one sample space outcome will occur. For example, if we consider the experiment of tossing a coin and observing whether the upward face of the coin shows as a "head" or a "tail," then the sample space consists of the outcomes "head" and "tail." If we consider the experiment of rolling a die and observing the number of dots showing on the upward face of the die, then the sample space consists of the outcomes 1, 2, 3, 4, 5, and 6. If we consider the experiment of subjecting an automobile to a "pass-fail" tailpipe emissions test, then the sample space consists of the outcomes "pass" and "fail."

Assigning probabilities to sample space outcomes We often wish to assign probabilities to sample space outcomes. This is usually done by using one of three methods: the *classical method*, the *relative frequency method,* or the *subjective method.* Regardless of the method used, **probabilities must be assigned to the experimental outcomes so that two conditions are met:**

1. The probability assigned to each sample space outcome must be between 0 and 1. That is, if E represents a sample space outcome and if $P(E)$ represents the probability of this outcome, then $0 \leq P(E) \leq 1$.
2. The probabilities of all of the sample space outcomes must sum to 1.

The **classical method** of assigning probabilities can be used when the sample space outcomes are equally likely. For example, consider the experiment of tossing a fair coin. Here, there are *two* equally likely sample space outcomes—head (H) and tail (T). Therefore, logic suggests that the probability of observing a head, denoted $P(H)$, is $1/2 = .5$, and that the probability of observing a tail, denoted $P(T)$, is also $1/2 = .5$. Notice that each probability is between 0 and 1. Furthermore, because H and T are all of the sample space outcomes, $P(H) + P(T) = 1$. In general, if there are N equally likely sample space outcomes, the probability assigned to each sample space outcome is $1/N$. To illustrate this, consider the experiment of rolling a fair die. It would seem reasonable to think that the six sample space outcomes 1, 2, 3, 4, 5, and 6 are equally likely, and thus each outcome is assigned a probability of $1/6$. If $P(1)$ denotes the probability that one dot appears on the upward face of the die, then $P(1) = 1/6$. Similarly, $P(2) = 1/6$, $P(3) = 1/6$, $P(4) = 1/6$, $P(5) = 1/6$, and $P(6) = 1/6$.

Before discussing the *relative frequency method* for assigning probabilities, we note that probability is often interpreted to be a **long run relative frequency.** To illustrate this, consider tossing a fair coin—a coin such that the probability of its upward face showing as a head is .5. If we get 6 heads in the first 10 tosses, then the relative frequency, or fraction, of heads is $6/10 = .6$. If we get 47 heads in the first 100 tosses, the relative frequency of heads is $47/100 = .47$. If we get 5,067 heads in the first 10,000 tosses, the relative frequency of heads is $5{,}067/10{,}000 = .5067$.[1] Note that the relative frequency of heads is approaching (that is, getting closer to) .5. The long run relative frequency interpretation of probability says that, if we tossed the coin an indefinitely large number of times (that is, a number of times *approaching infinity*), the relative frequency of heads obtained would approach .5. Of course, in actuality it is impossible to toss a coin (or perform any experiment) an indefinitely large number of times. Therefore, a relative frequency interpretation of probability is a mathematical idealization. To summarize, suppose that E is an experimental outcome that might occur when a particular experiment is performed. Then the probability that E will occur, $P(E)$, can be interpreted to be the number that would be approached by the relative frequency of E if we performed the experiment an indefinitely large number of times. It follows that we often think of a probability in terms of the percentage of the time the experimental outcome would occur in many repetitions of the experiment. For instance, when we say that the probability of obtaining a head when we toss a coin is .5, we are saying that, when we repeatedly toss the coin an indefinitely large number of times, we will obtain a head on 50 percent of the repetitions.

Sometimes it is either difficult or impossible to use the classical method to assign probabilities. Since we can often make a relative frequency interpretation of probability, we can estimate a probability by performing the experiment in which an outcome might occur many times. Then, we estimate the probability of the experimental outcome to be the proportion of the time that the outcome occurs during the many repetitions of the experiment. For example, to estimate the probability that a randomly selected consumer prefers Coca-Cola to all other soft drinks, we perform an experiment in which we ask a randomly selected consumer for his or her preference. There are two possible experimental outcomes: "prefers Coca-Cola" and "does not prefer Coca-Cola." However, we have no reason to believe that these experimental outcomes are equally likely, so we cannot use the classical method. We might perform the experiment, say, 1,000 times by surveying 1,000 randomly selected consumers. Then, if 140 of those surveyed said that they prefer Coca-Cola, we would estimate the probability that a randomly selected consumer prefers Coca-Cola to all other soft drinks to be $140/1{,}000 = .14$. This is an example of the **relative frequency method** of assigning probability.

If we cannot perform the experiment many times, we might estimate the probability by using our previous experience with similar situations, intuition, or special expertise that we may possess. For example, a company president might estimate the probability of success for a one-time business venture to be .7. Here, on the basis of knowledge of the success of previous similar ventures, the opinions of company personnel, and other pertinent information, the president believes that there is a 70 percent chance the venture will be successful.

[1]The South African mathematician John Kerrich actually obtained this result when he tossed a coin 10,000 times while imprisoned by the Germans during World War II.

When we use experience, intuitive judgement, or expertise to assess a probability, we call this the **subjective method** of assigning probability. Such a probability (called a **subjective probability**) may or may not have a relative frequency interpretation. For instance, when the company president estimates that the probability of a successful business venture is .7, this may mean that, if business conditions similar to those that are about to be encountered could be repeated many times, then the business venture would be successful in 70 percent of the repetitions. Or the president may not be thinking in relative frequency terms but rather may consider the venture a "one-shot" proposition. We will discuss some other subjective probabilities later. However, the interpretations of statistical inferences we will explain in later chapters are based on the relative frequency interpretation of probability. For this reason, we will concentrate on this interpretation.

4.2 Probability and Events

LO4-2 List the outcomes in a sample space and use the list to compute probabilities.

At the beginning of this chapter, we informally talked about events. We now give the formal definition of an event.

An **event** is a set of one or more sample space outcomes.

For example, if we consider the experiment of tossing a fair die, the event "at least five spots will show on the upward face of the die" consists of the sample space outcomes 5 and 6. That is, the event "at least five spots will show on the upward face of the die" will occur if and only if one of the sample space outcomes 5 or 6 occurs.

To find the probability that an event will occur, we can use the following result.

The **probability of an event** is the **sum of the probabilities of the sample space outcomes** that correspond to the event.

As an example, we have seen that if we consider the experiment of tossing a fair die, then the sample space outcomes 5 and 6 correspond to the occurrence of the event "at least five spots will show on the upward face of the die." Therefore, the probability of this event is

$$P(5) + P(6) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

EXAMPLE 4.1 Boys and Girls

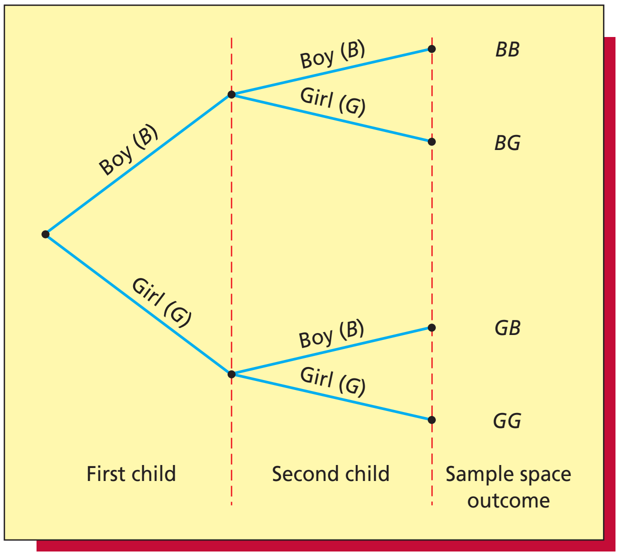
A newly married couple plans to have two children. Naturally, they are curious about whether their children will be boys or girls. Therefore, we consider the experiment of having two children. In order to find the sample space of this experiment, we let B denote that a child is a boy and G denote that a child is a girl. Then, it is useful to construct the tree diagram shown in Figure 4.1. This diagram pictures the experiment as a two-step process—having the first child, which could be either a boy or a girl (B or G), and then having the second child, which could also be either a boy or a girl (B or G). Each branch of the tree leads to a sample space outcome. These outcomes are listed at the right ends of the branches. We see that there are four sample space outcomes. Therefore, the sample space (that is, the set of all the sample space outcomes) is

$$BB \qquad BG \qquad GB \qquad GG$$

In order to consider the probabilities of these outcomes, suppose that boys and girls are equally likely each time a child is born. Intuitively, this says that each of the sample space outcomes is equally likely. That is, this implies that

$$P(BB) = P(BG) = P(GB) = P(GG) = \frac{1}{4}$$

FIGURE 4.1 A Tree Diagram of the Genders of Two Children



Therefore:

1. The probability that the couple will have two boys is

$$P(BB) = \frac{1}{4}$$

because two boys will be born if and only if the sample space outcome *BB* occurs.

2. The probability that the couple will have one boy and one girl is

$$P(BG) + P(GB) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

because one boy and one girl will be born if and only if one of the sample space outcomes *BG* or *GB* occurs.

3. The probability that the couple will have two girls is

$$P(GG) = \frac{1}{4}$$

because two girls will be born if and only if the sample space outcome *GG* occurs.

4. The probability that the couple will have at least one girl is

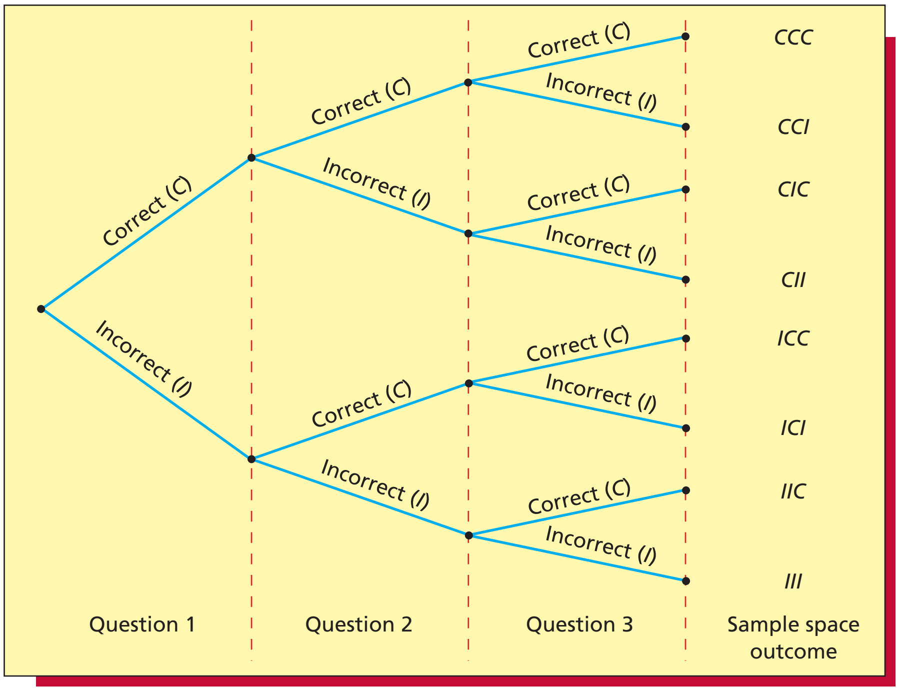
$$P(BG) + P(GB) + P(GG) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$$

because at least one girl will be born if and only if one of the sample space outcomes *BG*, *GB*, or *GG* occurs.

EXAMPLE 4.2 Pop Quizzes

A student takes a pop quiz that consists of three true–false questions. If we consider our experiment to be answering the three questions, each question can be answered correctly or incorrectly. We will let *C* denote answering a question correctly and *I* denote answering a

FIGURE 4.2 A Tree Diagram of Answering Three True–False Questions



question incorrectly. Figure 4.2 depicts a tree diagram of the sample space outcomes for the experiment. The diagram portrays the experiment as a three-step process—answering the first question (correctly or incorrectly, that is, C or I), answering the second question, and answering the third question. The tree diagram has eight different branches, and the eight sample space outcomes are listed at the ends of the branches. We see that the sample space is

CCC	*CCI*	*CIC*	*CII*
ICC	*ICI*	*IIC*	*III*

Next, suppose that the student was totally unprepared for the quiz and had to blindly guess the answer to each question. That is, the student had a 50–50 chance (or .5 probability) of correctly answering each question. Intuitively, this would say that each of the eight sample space outcomes is equally likely to occur. That is,

$$P(CCC) = P(CCI) = \cdots = P(III) = \frac{1}{8}$$

Therefore:

1 The probability that the student will get all three questions correct is

$$P(CCC) = \frac{1}{8}$$

2 The probability that the student will get exactly two questions correct is

$$P(CCI) + P(CIC) + P(ICC) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$$

because two questions will be answered correctly if and only if one of the sample space outcomes *CCI*, *CIC*, or *ICC* occurs.

3 The probability that the student will get exactly one question correct is

$$P(CII) + P(ICI) + P(IIC) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$$

because one question will be answered correctly if and only if one of the sample space outcomes *CII*, *ICI*, or *IIC* occurs.

4 The probability that the student will get all three questions incorrect is

$$P(III) = \frac{1}{8}$$

5 The probability that the student will get at least two questions correct is

$$P(CCC) + P(CCI) + P(CIC) + P(ICC) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{1}{2}$$

because the student will get at least two questions correct if and only if one of the sample space outcomes *CCC*, *CCI*, *CIC*, or *ICC* occurs.

Notice that in the true–false question situation, we find that, for instance, the probability that the student will get exactly one question correct equals the ratio

$$\frac{\text{the number of sample space outcomes resulting in one correct answer}}{\text{the total number of sample space outcomes}} = \frac{3}{8}$$

In general, when a sample space is finite we can use the following method for computing the probability of an event.

If all of the sample space outcomes are equally likely, then the probability that an event will occur is equal to the ratio

$$\frac{\text{the number of sample space outcomes that correspond to the event}}{\text{the total number of sample space outcomes}}$$

It is important to emphasize, however, that we can use this rule only when all of the sample space outcomes are equally likely (as they are in the true–false question situation). If the sample space outcomes are not equally likely, the rule may give an incorrect probability.

EXAMPLE 4.3 Choosing a CEO

A company is choosing a new chief executive officer (CEO). It has narrowed the list of candidates to four finalists (identified by last name only)—Adams, Chung, Hill, and Rankin. If we consider our experiment to be making a final choice of the company's CEO, then the experiment's sample space consists of the four possible outcomes:

$A \equiv$ Adams will be chosen as CEO.

$C \equiv$ Chung will be chosen as CEO.

$H \equiv$ Hill will be chosen as CEO.

$R \equiv$ Rankin will be chosen as CEO.

Next, suppose that industry analysts feel (subjectively) that the probabilities that Adams, Chung, Hill, and Rankin will be chosen as CEO are .1, .2, .5, and .2, respectively. That is, in probability notation

$$P(A) = .1 \qquad P(C) = .2 \qquad P(H) = .5 \qquad \text{and} \qquad P(R) = .2$$

Also, suppose only Adams and Hill are internal candidates (they already work for the company). Letting *INT* denote the event that "an internal candidate will be selected for the CEO position," then *INT* consists of the sample space outcomes *A* and *H* (that is, *INT* will occur if and only if either of the sample space outcomes *A* or *H* occurs). It follows that $P(INT) = P(A) + P(H) = .1 + .5 = .6$. This says that the probability that an internal candidate will be chosen to be CEO is .6.

Finally, it is important to understand that if we had ignored the fact that sample space outcomes are not equally likely, we might have tried to calculate $P(INT)$ as follows:

$$P(INT) = \frac{\text{the number of internal candidates}}{\text{the total number of candidates}} = \frac{2}{4} = .5$$

This result would be incorrect. Because the sample space outcomes are not equally likely, we have seen that the correct value of $P(INT)$ is .6, not .5.

EXAMPLE 4.4 The Crystal Cable Case: Market Penetration

Like all companies, cable companies send shareholders reports on their profits, dividends, and return on equity. They often supplement this information with some metrics unique to the cable business. To construct one such metric, a cable company can compare the number of households it actually serves to the number of households its current transmission lines could reach (without extending the lines). The number of households that the cable company's lines could reach is called its number of **cable passings,** while the ratio of the number of households the cable company actually serves to its number of cable passings is called the company's **cable penetration.** There are various types of cable penetrations—one for cable television, one for cable Internet, one for cable phone, and others. Moreover, a cable penetration is a probability, and interpreting it as such will help us to better understand various techniques to be discussed in the next section. For example, in a recent quarterly report, Crystal Cable reported that it had 12.4 million cable television customers and 27.4 million cable passings.[2] Consider randomly selecting one of Crystal's cable passings. That is, consider selecting one cable passing by giving each and every cable passing the same chance of being selected. Let *A* be the event that the randomly selected cable passing has Crystal's cable television service. Then, because the sample space of this experiment consists of 27.4 million equally likely sample space outcomes (cable passings), it follows that

$$\begin{aligned} P(A) &= \frac{\text{the number of cable passings that have Crystal's cable television service}}{\text{the total number of cable passings}} \\ &= \frac{12.4}{27.4} \\ &= .45 \end{aligned}$$

This probability is Crystal's cable television penetration and says that the probability that a randomly selected cable passing has Crystal's cable television service is .45. That is, 45 percent of Crystal's cable passings have Crystal's cable television service.

To conclude this section, we note that in optional Section 4.6 we discuss several *counting rules* that can be used to count the number of sample space outcomes in an experiment. These rules are particularly useful when there are many sample space outcomes and thus these outcomes are difficult to list.

[2]Although these numbers are hypothetical, they are similar to results actually found in Time Warner Cable's quarterly reports. See www.TimeWarnerCable.com. Click on Investor Relations.

Exercises for Sections 4.1 and 4.2

CONCEPTS

4.1 Define the following terms: *experiment, event, probability, sample space.*

4.2 Explain the properties that must be satisfied by a probability.

METHODS AND APPLICATIONS

4.3 Two randomly selected grocery store patrons are each asked to take a blind taste test and to then state which of three diet colas (marked as A, B, or C) he or she prefers.

a Draw a tree diagram depicting the sample space outcomes for the test results.

b List the sample space outcomes that correspond to each of the following events:

(1) Both patrons prefer diet cola A.

(2) The two patrons prefer the same diet cola.

(3) The two patrons prefer different diet colas.

(4) Diet cola A is preferred by at least one of the two patrons.

(5) Neither of the patrons prefers diet cola C.

c Assuming that all sample space outcomes are equally likely, find the probability of each of the events given in part *b*.

4.4 Suppose that a couple will have three children. Letting B denote a boy and G denote a girl:

a Draw a tree diagram depicting the sample space outcomes for this experiment.

b List the sample space outcomes that correspond to each of the following events:

(1) All three children will have the same gender.

(2) Exactly two of the three children will be girls.

(3) Exactly one of the three children will be a girl.

(4) None of the three children will be a girl.

c Assuming that all sample space outcomes are equally likely, find the probability of each of the events given in part *b*.

4.5 Four people will enter an automobile showroom, and each will either purchase a car (P) or not purchase a car (N).

a Draw a tree diagram depicting the sample space of all possible purchase decisions that could potentially be made by the four people.

b List the sample space outcomes that correspond to each of the following events:

(1) Exactly three people will purchase a car.

(2) Two or fewer people will purchase a car.

(3) One or more people will purchase a car.

(4) All four people will make the same purchase decision.

c Assuming that all sample space outcomes are equally likely, find the probability of each of the events given in part *b*.

4.6 The U.S. Census Bureau compiles data on family income and summarizes its findings in *Current Population Reports.* The table below is a frequency distribution of the annual incomes for a random sample of U.S. families. Find an estimate of the probability that a randomly selected U.S. family has an income between \$60,000 and \$199,999.

Income	Frequency (in thousands)
Under \$20,000	11,470
\$20,000–\$39,999	17,572
\$40,000–\$59,999	14,534
\$60,000–\$79,999	11,410
\$80,000–\$99,999	7,535
\$100,000–\$199,999	11,197
\$200,000 and above	2,280
	75,998

4.7 Let A, B, C, D, and E be sample space outcomes forming a sample space. Suppose that $P(A) = .2$, $P(B) = .15$, $P(C) = .3$, and $P(D) = .2$. What is $P(E)$? Explain how you got your answer.

4.3 Some Elementary Probability Rules

LO4-3 Use elementary probability rules to compute probabilities.

We can often calculate probabilities by using formulas called **probability rules.** We will begin by presenting the simplest probability rule: the *rule of complements.* To start, we define the complement of an event:

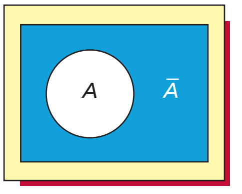
Given an event A, the **complement of A** is the event consisting of all sample space outcomes that do not correspond to the occurrence of A. The complement of A is denoted $\overline{A}$. Furthermore, $P(\overline{A})$ denotes **the probability that A will not occur.**

Figure 4.3 is a **Venn diagram** depicting the complement $\overline{A}$ of an event A. In any probability situation, either an event A or its complement $\overline{A}$ must occur. Therefore, we have

$$P(A) + P(\overline{A}) = 1$$

This implies the following result:

FIGURE 4.3 The Complement of an Event (the Shaded Region Is $\overline{A}$, the Complement of A)



The Rule of Complements

Consider an event A. Then, **the probability that A *will not occur*** is

$$P(\overline{A}) = 1 - P(A)$$

EXAMPLE 4.5 The Crystal Cable Case: Market Penetration

Recall from Example 4.4 that the probability that a randomly selected cable passing has Crystal's cable television service is .45. It follows that the probability of the complement of this event (that is, the probability that a randomly selected cable passing does not have Crystal's cable television service) is $1 - .45 = .55$.

We next define the *intersection* of two events. Consider performing an experiment a single time. Then:

Given two events A and B, the **intersection of A and B** is the event that occurs if both A and B simultaneously occur. The intersection is denoted by $A \cap B$. Furthermore, $P(A \cap B)$ denotes **the probability that *both* A *and* B *will simultaneously occur.***

EXAMPLE 4.6 The Crystal Cable Case: Market Penetration

Recall from Example 4.4 that Crystal Cable has 27.4 million cable passings. Consider randomly selecting one of these cable passings, and define the following events:

$A \equiv$ the randomly selected cable passing has Crystal's cable television service.

$\overline{A} \equiv$ the randomly selected cable passing does not have Crystal's cable television service.

$B \equiv$ the randomly selected cable passing has Crystal's cable Internet service.

$\overline{B} \equiv$ the randomly selected cable passing does not have Crystal's cable Internet service.

$A \cap B \equiv$ the randomly selected cable passing has both Crystal's cable television service and Crystal's cable Internet service.

$A \cap \overline{B} \equiv$ the randomly selected cable passing has Crystal's cable television service and does not have Crystal's cable Internet service.

TABLE 4.1 A Contingency Table Summarizing Crystal's Cable Television and Internet Penetration (Figures In Millions Of Cable Passings)

Events	Has Cable Internet Service, B	Does Not Have Cable Internet Service, $\overline{B}$	Total
Has Cable Television Service, A	6.5	5.9	12.4
Does Not Have Cable Television Service, $\overline{A}$	3.3	11.7	15.0
Total	9.8	17.6	27.4

$\overline{A} \cap B \equiv$ the randomly selected cable passing does not have Crystal's cable television service and does have Crystal's cable Internet service.

$\overline{A} \cap \overline{B} \equiv$ the randomly selected cable passing does not have Crystal's cable television service and does not have Crystal's cable Internet service.

Table 4.1 is a *contingency table* that summarizes Crystal's cable passings. Using this table, we can calculate the following probabilities, each of which describes some aspect of Crystal's cable penetrations:

1. Because 12.4 million out of 27.4 million cable passings have Crystal's cable television service, A, then

$$P(A) = \frac{12.4}{27.4} = .45$$

This says that 45 percent of Crystal's cable passings have Crystal's cable television service (as previously seen in Example 4.4).

2. Because 9.8 million out of 27.4 million cable passings have Crystal's cable Internet service, B, then

$$P(B) = \frac{9.8}{27.4} = .36$$

This says that 36 percent of Crystal's cable passings have Crystal's cable Internet service.

3. Because 6.5 million out of 27.4 million cable passings have Crystal's cable television service and Crystal's cable Internet service, $A \cap B$, then

$$P(A \cap B) = \frac{6.5}{27.4} = .24$$

This says that 24 percent of Crystal's cable passings have both of Crystal's cable services.

4. Because 5.9 million out of 27.4 million cable passings have Crystal's cable television service, but do not have Crystal's cable Internet service, $A \cap \overline{B}$, then

$$P(A \cap \overline{B}) = \frac{5.9}{27.4} = .22$$

This says that 22 percent of Crystal's cable passings have only Crystal's cable television service.

5 Because 3.3 million out of 27.4 million cable passings do not have Crystal's cable television service, but do have Crystal's cable Internet service, $\overline{A} \cap B$, then

$$P(\overline{A} \cap B) = \frac{3.3}{27.4} = .12$$

This says that 12 percent of Crystal's cable passings have only Crystal's cable Internet service.

6 Because 11.7 million out of 27.4 million cable passings do not have Crystal's cable television service and do not have Crystal's cable Internet service, $\overline{A} \cap \overline{B}$, then

$$P(\overline{A} \cap \overline{B}) = \frac{11.7}{27.4} = .43$$

This says that 43 percent of Crystal's cable passings have neither of Crystal's cable services.

We next consider the *union* of two events. Again consider performing an experiment a single time. Then:

Given two events A and B, the **union of A and B** is the event that occurs if A or B (or both) occur. The union is denoted $A \cup B$. Furthermore, $P(A \cup B)$ denotes **the probability that *A or B (or both) will occur.***

EXAMPLE 4.7 The Crystal Cable Case: Market Penetration

Consider randomly selecting one of Crystal's 27.4 million cable passings, and define the event

$A \cup B \equiv$ the randomly selected cable passing has Crystal's cable television service or Crystal's cable Internet service (or both)—that is, has at least one of the two services.

Looking at Table 4.1, we see that the cable passings that have Crystal's cable television service or Crystal's cable Internet service are (1) the 5.9 million cable passings that have only Crystal's cable television service, $A \cap \overline{B}$, (2) the 3.3 million cable passings that have only Crystal's cable Internet service, $\overline{A} \cap B$, and (3) the 6.5 million cable passings that have both Crystal's cable television service and Crystal's cable Internet service, $A \cap B$. Therefore, because a total of 15.7 million cable passings have Crystal's cable television service or Crystal's cable Internet service (or both), it follows that

$$P(A \cup B) = \frac{15.7}{27.4} = .57$$

This says that the probability that the randomly selected cable passing has Crystal's cable television service or Crystal's cable Internet service (or both) is .57. That is, 57 percent of Crystal's cable passings have Crystal's cable television service or Crystal's cable Internet service (or both). Notice that $P(A \cup B) = .57$ does not equal

$$P(A) + P(B) = .45 + .36 = .81$$

Logically, the reason for this is that both $P(A) = .45$ and $P(B) = .36$ count the 24 percent of the cable passings that have both Crystal's cable television service and Crystal's cable Internet service. Therefore, the sum of $P(A)$ and $P(B)$ counts this 24 percent of the cable passings once too often. It follows that if we subtract $P(A \cap B) = .24$ from the sum of $P(A)$ and $P(B)$, then we will obtain $P(A \cup B)$. That is,

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= .45 + .36 - .24 = .57 \end{aligned}$$

Noting that Figure 4.4 shows **Venn diagrams** depicting the events A, B, $A \cap B$, and $A \cup B$, we have the following general result:

The Addition Rule

Let A and B be events. Then, **the probability that *A or B (or both) will occur*** is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

The reasoning behind this result has been illustrated at the end of Example 4.7. Similarly, the Venn diagrams in Figure 4.4 show that when we compute $P(A) + P(B)$, we are counting each of the sample space outcomes in $A \cap B$ twice. We correct for this by subtracting $P(A \cap B)$.

We next define the idea of *mutually exclusive events:*

Mutually Exclusive Events

Two events A and B are **mutually exclusive** if they have no sample space outcomes in common. In this case, the events A and B cannot occur simultaneously, and thus

$$P(A \cap B) = 0$$

Noting that Figure 4.5 is a Venn diagram depicting two mutually exclusive events, we consider the following example.

EXAMPLE 4.8 Selecting Playing Cards

Consider randomly selecting a card from a standard deck of 52 playing cards. We define the following events:

$J \equiv$ the randomly selected card is a jack.

$Q \equiv$ the randomly selected card is a queen.

$R \equiv$ the randomly selected card is a red card (that is, a diamond or a heart).

Because there is no card that is both a jack and a queen, the events J and Q are mutually exclusive. On the other hand, there are two cards that are both jacks and red cards—the jack of diamonds and the jack of hearts—so the events J and R are not mutually exclusive.

We have seen that for any two events A and B, the probability that A or B (or both) will occur is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Therefore, when calculating $P(A \cup B)$, we should always subtract $P(A \cap B)$ from the sum of $P(A)$ and $P(B)$. However, when A and B are mutually exclusive, $P(A \cap B)$ equals 0. Therefore, in this case—and only in this case—we have the following:

The Addition Rule for Two Mutually Exclusive Events

Let A and B be **mutually exclusive** events. Then, **the probability that *A or B will occur*** is

$$P(A \cup B) = P(A) + P(B)$$

FIGURE 4.4 Venn Diagrams Depicting the Events A, B, $A \cap B$, and $A \cup B$

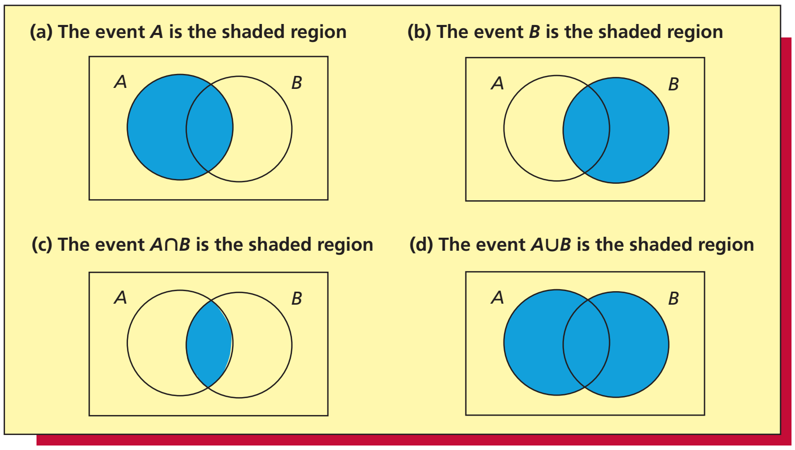
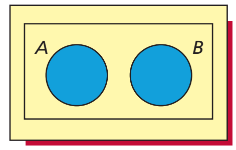


FIGURE 4.5 Two Mutually Exclusive Events



EXAMPLE 4.9 Selecting Playing Cards

Again consider randomly selecting a card from a standard deck of 52 playing cards, and define the events

$J \equiv$ the randomly selected card is a jack.

$Q \equiv$ the randomly selected card is a queen.

$R \equiv$ the randomly selected card is a red card (a diamond or a heart).

Because there are four jacks, four queens, and 26 red cards, we have $P(J) = \frac{4}{52}$, $P(Q) = \frac{4}{52}$, and $P(R) = \frac{26}{52}$. Furthermore, since there is no card that is both a jack and a queen, the events J and Q are mutually exclusive and thus $P(J \cap Q) = 0$. It follows that the probability that the randomly selected card is a jack or a queen is

$$\begin{aligned} P(J \cup Q) &= P(J) + P(Q) \\ &= \frac{4}{52} + \frac{4}{52} = \frac{8}{52} = \frac{2}{13} \end{aligned}$$

Because there are two cards that are both jacks and red cards—the jack of diamonds and the jack of hearts—the events J and R are not mutually exclusive. Therefore, the probability that the randomly selected card is a jack or a red card is

$$\begin{aligned} P(J \cup R) &= P(J) + P(R) - P(J \cap R) \\ &= \frac{4}{52} + \frac{26}{52} - \frac{2}{52} = \frac{28}{52} = \frac{7}{13} \end{aligned}$$

We now consider an arbitrary group of events—$A_1, A_2, \ldots, A_N$. We will denote the probability that A_1 or A_2 or $\cdots$ or A_N occurs (that is, the probability that at least one of the events occurs) as $P(A_1 \cup A_2 \cup \cdots \cup A_N)$. Although there is a formula for this probability, it is quite complicated and we will not present it in this book. However, sometimes we can use sample spaces to reason out such a probability. For instance, in the playing card situation of Example 4.9, there are four jacks, four queens, and 22 red cards that are not jacks or queens (the 26 red cards minus the two red jacks and the two red queens). Therefore, because there are a total of 30 cards corresponding to the event $J \cup Q \cup R$, it follows that

$$P(J \cup Q \cup R) = \frac{30}{52} = \frac{15}{26}$$

Because some cards are both jacks and red cards, and because some cards are both queens and red cards, we say that the events J, Q, and R are not mutually exclusive. When, however, a group of events is mutually exclusive, there is a simple formula for the probability that at least one of the events will occur:

The Addition Rule for N Mutually Exclusive Events

The events $A_1, A_2, \ldots, A_N$ are mutually exclusive if no two of the events have any sample space outcomes in common. In this case, no two of the events can occur simultaneously, and

$$P(A_1 \cup A_2 \cup \cdots \cup A_N) = P(A_1) + P(A_2) + \cdots + P(A_N)$$

As an example of using this formula, again consider the playing card situation and the events J and Q. If we define the event

$$K \equiv \text{the randomly selected card is a king}$$

then the events J, Q, and K are mutually exclusive. Therefore,

$$\begin{aligned} P(J \cup Q \cup K) &= P(J) + P(Q) + P(K) \\ &= \frac{4}{52} + \frac{4}{52} + \frac{4}{52} = \frac{12}{52} = \frac{3}{13} \end{aligned}$$

Exercises for Section 4.3

CONCEPTS

connect

4.8 Explain what it means for two events to be mutually exclusive; for N events.

4.9 If A and B are events, define (in words) $\overline{A}$, $A \cup B$, $A \cap B$, and $\overline{A} \cap \overline{B}$.

METHODS AND APPLICATIONS

4.10 Consider a standard deck of 52 playing cards, a randomly selected card from the deck, and the following events:

R = red B = black A = ace N = nine D = diamond C = club

a Describe the sample space outcomes that correspond to each of these events.

b For each of the following pairs of events, indicate whether the events are mutually exclusive. In each case, if you think the events are mutually exclusive, explain why the events have no common sample space outcomes. If you think the events are not mutually exclusive, list the sample space outcomes that are common to both events.

(1) R and A (3) A and N (5) D and C
(2) R and C (4) N and C

4.11 The following contingency table summarizes the number of students at a college who have a Mastercard or a Visa credit card.

	Have Visa	Do Not Have Visa	Total
Have Mastercard	1,000	1,500	2,500
Do not have Mastercard	3,000	4,500	7,500
Total	4,000	6,000	10,000

a Find the probability that a randomly selected student
(1) Has a Mastercard.
(2) Has a VISA.
(3) Has both credit cards.

TABLE 4.2 Results of a Concept Study for a New Wine Cooler DS WineCooler

Rating	Total	Gender Male	Gender Female	Age Group 21–24	Age Group 25–34	Age Group 35–49
Extremely appealing (5)	151	68	83	48	66	37
(4)	91	51	40	36	36	19
(3)	36	21	15	9	12	15
(2)	13	7	6	4	6	3
Not at all appealing (1)	9	3	6	4	3	2

Source: W. R. Dillon, T. J. Madden, and N. H. Firtle, *Essentials of Marketing Research* (Burr Ridge, IL: Richard D. Irwin, Inc., 1993), p. 390.

b Find the probability that a randomly selected student
 (1) Has a Mastercard or a VISA.
 (2) Has neither credit card.
 (3) Has exactly one of the two credit cards.

4.12 The card game of Euchre employs a deck that consists of all four of each of the aces, kings, queens, jacks, tens, and nines (one of each suit—clubs, diamonds, spades, and hearts). Find the probability that a randomly selected card from a Euchre deck is a jack (J), a spade (S), a jack or an ace (A), a jack or a spade. Are the events J and A mutually exclusive? J and S? Why or why not?

4.13 Each month a brokerage house studies various companies and rates each company's stock as being either "low risk" or "moderate to high risk." In a recent report, the brokerage house summarized its findings about 15 aerospace companies and 25 food retailers in the following table:

Company Type	Low Risk	Moderate to High Risk
Aerospace company	6	9
Food retailer	15	10

If we randomly select one of the total of 40 companies, find
a The probability that the company is a food retailer.
b The probability that the company's stock is "low risk."
c The probability that the company's stock is "moderate to high risk."
d The probability that the company is a food retailer and has a stock that is "low risk."
e The probability that the company is a food retailer or has a stock that is "low risk."

4.14 In the book *Essentials of Marketing Research,* William R. Dillon, Thomas J. Madden, and Neil H. Firtle present the results of a concept study for a new wine cooler. Three hundred consumers between 21 and 49 years old were randomly selected. After sampling the new beverage, each was asked to rate the appeal of the phrase

> Not sweet like wine coolers, not filling like beer, and more refreshing than wine or mixed drinks

as it relates to the new wine cooler. The rating was made on a scale from 1 to 5, with 5 representing "extremely appealing" and with 1 representing "not at all appealing." The results obtained are given in Table 4.2. Estimate the probability that a randomly selected 21- to 49-year-old consumer
a Would give the phrase a rating of 5.
b Would give the phrase a rating of 3 or higher.
c Is in the 21–24 age group; the 25–34 age group; the 35–49 age group.
d Is a male who gives the phrase a rating of 4.
e Is a 35- to 49-year-old who gives the phrase a rating of 1.

4.15 In Exercise 4.14 estimate the probability that a randomly selected 21- to 49-year-old consumer is a 25- to 49-year-old who gives the phrase a rating of 5.

4.4 Conditional Probability and Independence

LO4-4 Compute conditional probabilities and assess independence.

Conditional probability In Table 4.3 we repeat Table 4.1 summarizing data concerning Crystal cable's 27.4 million cable passings. Suppose that we randomly select a cable passing and that the chosen cable passing reports that it has Crystal's cable Internet service. Given this new

TABLE 4.3 A Contingency Table Summarizing Crystal's Cable Television and Internet Penetration (Figures In Millions Of Cable Passings)

Events	Has Cable Internet Service, B	Does Not Have Cable Internet Service, $\overline{B}$	Total
Has Cable Television Service, A	6.5	5.9	12.4
Does Not Have Cable Television Service, $\overline{A}$	3.3	11.7	15.0
Total	9.8	17.6	27.4

information, we wish to find the probability that the cable passing has Crystal's cable television service. This new probability is called a **conditional probability.**

The **probability of the event A, given the condition that the event B *has occurred,*** is written as $P(A|B)$—pronounced "the probability of A given B." We often refer to such a probability as the **conditional probability of A given B.**

In order to find the conditional probability that a randomly selected cable passing has Crystal's cable television service, given that it has Crystal's cable Internet service, notice that if we know that the randomly selected cable passing has Crystal's cable Internet service, we know that we are considering one of Crystal's 9.8 million cable Internet customers (see Table 4.3). That is, we are now considering what we might call a **reduced sample space** of Crystal's 9.8 million cable Internet customers. Because 6.5 million of these 9.8 million cable Internet customers also have Crystal's cable television service, we have

$$P(A|B) = \frac{6.5}{9.8} = .66$$

This says that the probability that the randomly selected cable passing has Crystal's cable television service, given that it has Crystal's cable Internet service, is .66. That is, 66 percent of Crystal's cable Internet customers also have Crystal's cable television service.

Next, suppose that we randomly select another cable passing from Crystal's 27.4 million cable passings, and suppose that this newly chosen cable passing reports that it has Crystal's cable television service. We now wish to find the probability that this cable passing has Crystal's cable Internet service. We write this new probability as $P(B|A)$. If we know that the randomly selected cable passing has Crystal's cable television service, we know that we are considering a reduced sample space of Crystal's 12.4 million cable television customers (see Table 4.3). Because 6.5 million of these 12.4 million cable television customers also have Crystal's cable Internet service, we have

$$P(B|A) = \frac{6.5}{12.4} = .52$$

This says that the probability that the randomly selected cable passing has Crystal's cable Internet service, given that it has Crystal's cable television service, is .52. That is, 52 percent of Crystal's cable television customers also have Crystal's cable Internet service.

If we divide both the numerator and denominator of each of the conditional probabilities $P(A \mid B)$ and $P(B \mid A)$ by 27.4, we obtain

$$P(A|B) = \frac{6.5}{9.8} = \frac{6.5/27.4}{9.8/27.4} = \frac{P(A \cap B)}{P(B)}$$

$$P(B|A) = \frac{6.5}{12.4} = \frac{6.5/27.4}{12.4/27.4} = \frac{P(A \cap B)}{P(A)}$$

We express these conditional probabilities in terms of $P(A)$, $P(B)$, and $P(A \cap B)$ in order to obtain a more general formula for a conditional probability. We need a more general formula because, although we can use the reduced sample space approach we have demonstrated to find conditional probabilities when all of the sample space outcomes are equally likely, this approach may not give correct results when the sample space outcomes are *not* equally likely. We now give expressions for conditional probability that are valid for any sample space.

Conditional Probability

1 The **conditional probability of the event *A* given that the event *B* has occurred** is written ***P*(*A* | *B*)** and is defined to be

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Here we assume that $P(B)$ is greater than 0.

2 The **conditional probability of the event *B* given that the event *A* has occurred** is written ***P*(*B* | *A*)** and is defined to be

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Here we assume that $P(A)$ is greater than 0.

If we multiply both sides of the equation

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

by $P(B)$, we obtain the equation

$$P(A \cap B) = P(B)P(A|B)$$

Similarly, if we multiply both sides of the equation

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

by $P(A)$, we obtain the equation

$$P(A \cap B) = P(A)P(B|A)$$

In summary, we now have two equations that can be used to calculate $P(A \cap B)$. These equations are often referred to as the **general multiplication rule** for probabilities.

The General Multiplication Rule—Two Ways to Calculate $P(A \cap B)$

Given any two events A and B,

$$\begin{aligned} P(A \cap B) &= P(A)P(B|A) \\ &= P(B)P(A|B) \end{aligned}$$

EXAMPLE 4.10 Gender Issues at a Pharmaceutical Company

C

At a large pharmaceutical company, 52 percent of the sales representatives are women, and 44 percent of the sales representatives having a management position are women. (There are various types of management positions in the sales division of the pharmaceutical company.) Given that 25 percent of the sales representatives have a management position, we wish to find

- The percentage of the sales representatives that have a management position and are women.
- The percentage of the female sales representatives that have a management position.
- The percentage of the sales representatives that have a management position and are men.
- The percentage of the male sales representatives that have a management position.

In order to find these percentages, consider randomly selecting one of the sales representatives. Then, let W denote the event that the randomly selected sales representative is a woman, and let M denote the event that the randomly selected sales representative is a man. Also, let MGT denote the event that the randomly selected sales representative has a management position. The information given at the beginning of this example says that 52 percent of the sales representatives are women and 44 percent of the sales representatives having a management position are women. This implies that $P(W) = .52$ and that $P(W \mid MGT) = .44$. The information given at the beginning of this example also says that 25 percent of the sales representatives have a management position. This implies that $P(MGT) = .25$. To find the percentage of the sales representatives that have a management position and are women, we find $P(MGT \cap W)$. The general multiplication rule tells us that

$$P(MGT \cap W) = P(MGT)P(W \mid MGT) = P(W)P(MGT \mid W)$$

Although we know that $P(W) = .52$, we do not know $P(MGT \mid W)$. Therefore, we cannot calculate $P(MGT \cap W)$ as $P(W)P(MGT \mid W)$. However, because we know that $P(MGT) = .25$ and $P(W \mid MGT) = .44$, we can calculate

$$P(MGT \cap W) = P(MGT)P(W \mid MGT) = (.25)(.44) = .11$$

This says that 11 percent of the sales representatives have a management position and are women. Moreover,

$$P(MGT \mid W) = \frac{P(MGT \cap W)}{P(W)} = \frac{.11}{.52} = .2115$$

This says that 21.15 percent of the female sales representatives have a management position.

To find the percentage of the sales representatives that have a management position and are men, we find $P(MGT \cap M)$. Because we know that 52 percent of the sales representatives are women, the rule of complements tells us that 48 percent of the sales representatives are men. That is, $P(M) = .48$. We also know that 44 percent of the sales representatives having a management position are women. It follows (by an extension of the rule of complements) that 56 percent of the sales representatives having a management position are men. That is, $P(M \mid MGT) = .56$. Using the fact that $P(MGT) = .25$, the general multiplication rule implies that

$$P(MGT \cap M) = P(MGT)P(M \mid MGT) = (.25)(.56) = .14$$

This says that 14 percent of the sales representatives have a management position and are men. Moreover,

$$P(MGT \mid M) = \frac{P(MGT \cap M)}{P(M)} = \frac{.14}{.48} = .2917$$

This says that 29.17 percent of the male sales representatives have a management position.

We have seen that $P(MGT) = .25$, while $P(MGT \mid W) = .2115$. Because $P(MGT \mid W)$ is less than $P(MGT)$, the probability that a randomly selected sales representative will have a management position is smaller if we know that the sales representative is a woman than it is if we have no knowledge of the sales representative's gender. Another way to see this is to recall that $P(MGT \mid M) = .2917$. Because $P(MGT \mid W) = .2115$ is less than $P(MGT \mid M) = .2917$, the probability that a randomly selected sales representative will have a management position is smaller if the sales representative is a woman than it is if the sales representative is a man.

Independence In Example 4.10 the probability of the event *MGT* is influenced by whether the event *W* occurs. In such a case, we say that the events *MGT* and *W* are **dependent.** If $P(MGT \mid W)$ were equal to $P(MGT)$, then the probability of the event *MGT* would not be influenced by whether *W* occurs. In this case we would say that the events *MGT* and *W* are **independent.** This leads to the following definition:

Independent Events

Two events *A* and *B* are **independent** if and only if

1. $P(A \mid B) = P(A)$ or, equivalently,
2. $P(B \mid A) = P(B)$

Here we assume that $P(A)$ and $P(B)$ are greater than 0.

EXAMPLE 4.11 Gender Issues at a Pharmaceutical Company

C

Recall that 52 percent of the pharmaceutical company's sales representatives are women. If 52 percent of the sales representatives having a management position were also women, then $P(W \mid MGT)$ would equal $P(W) = .52$. Moreover, recalling that $P(MGT) = .25$, it would follow that

$$P(MGT \mid W) = \frac{P(MGT \cap W)}{P(W)} = \frac{P(MGT)P(W \mid MGT)}{P(W)} = \frac{P(MGT)(.52)}{.52} = P(MGT)$$

That is, 25 percent of the female sales representatives—as well as 25 percent of all of the sales representatives—would have a management position. Of course, this independence is only hypothetical. The actual pharmaceutical company data led us to conclude that *MGT* and *W* are dependent. Specifically, because $P(MGT \mid W) = .2115$ is less than $P(MGT) = .25$ and $P(MGT \mid M) = .2917$, we conclude that women are less likely to have a management position at the pharmaceutical company. Looking at this another way, note that the ratio of $P(MGT \mid M) = .2917$ to $P(MGT \mid W) = .2115$ is $.2917/.2115 = 1.3792$. This says that **the probability that a randomly selected sales representative will have a management position is 37.92 percent higher if the sales representative is a man than it is if the sales representative is a woman.** Moreover, this conclusion describes the actual employment conditions that existed at Novartis Pharmaceutical Company from 2002 to 2007.[3] In the largest gender discrimination case ever to go to trial, Sanford, Wittels, and Heisler LLP used data implying the conclusion above—along with evidence of salary inequities and women being subjected to a hostile and sexist work environment—to successfully represent a class of 5,600 female sales representatives against Novartis. On May 19, 2010, a federal jury awarded $250 million to the class. The award was the largest ever in an employment discrimination case, and in November 2010 a final settlement agreement between Novartis and its female sales representatives was reached.

If the occurrences of the events *A* and *B* have nothing to do with each other, then we know that *A* and *B* are independent events. This implies that $P(A \mid B)$ equals $P(A)$ and that $P(B \mid A)$ equals $P(B)$. Recall that the general multiplication rule tells us that, for any two events *A* and *B*, we can say that $P(A \cap B) = P(A)P(B \mid A)$. Therefore, if $P(B \mid A)$ equals $P(B)$, it follows that $P(A \cap B) = P(A)P(B)$.

[3]Source: http://www.bononilawgroup.com/blog/2010/07/women-win-a-bias-suit-against-novartis.shtml

This equation is called the **multiplication rule for independent events.** To summarize:

The Multiplication Rule for Two Independent Events

If A and B are **independent events,** then

$$P(A \cap B) = P(A)P(B)$$

As a simple example, let C denote the event that your favorite college football team wins its first game next season, and let P denote the event that your favorite professional football team wins its first game next season. Suppose you believe that $P(C) = .6$ and $P(P) = .6$. Then, because the outcomes of a college football game and a professional football game would probably have nothing to do with each other, it is reasonable to assume that C and P are independent events. It follows that

$$P(C \cap P) = P(C)P(P) = (.6)(.6) = .36$$

This probability might seem surprisingly low. That is, because you believe that each of your teams has a 60 percent chance of winning, you might feel reasonably confident that both your college and professional teams will win their first game. Yet the chance of this happening is really only .36!

Next, consider a group of events $A_1, A_2, \ldots, A_N$. Intuitively, the events $A_1, A_2, \ldots, A_N$ are independent if the occurrences of these events have nothing to do with each other. Denoting the probability that all of these events will simultaneously occur as $P(A_1 \cap A_2 \cap \cdots \cap A_N)$, we have the following:

The Multiplication Rule for N Independent Events

If $A_1, A_2, \ldots, A_N$ are independent events, then

$$P(A_1 \cap A_2 \cap \cdots \cap A_N) = P(A_1)P(A_2) \cdots P(A_N)$$

EXAMPLE 4.12 An Application of the Independence Rule: Customer Service C

This example is based on a real situation encountered by a major producer and marketer of consumer products. The company assessed the service it provides by surveying the attitudes of its customers regarding 10 different aspects of customer service—order filled correctly, billing amount on invoice correct, delivery made on time, and so forth. When the survey results were analyzed, the company was dismayed to learn that only 59 percent of the survey participants indicated that they were satisfied with all 10 aspects of the company's service. Upon investigation, each of the 10 departments responsible for the aspects of service considered in the study insisted that it satisfied its customers 95 percent of the time. That is, each department claimed that its error rate was only 5 percent. Company executives were confused and felt that there was a substantial discrepancy between the survey results and the claims of the departments providing the services. However, a company statistician pointed out that there was no discrepancy. To understand this, consider randomly selecting a customer from among the survey participants, and define 10 events (corresponding to the 10 aspects of service studied):

$A_1 \equiv$ the customer is satisfied that the order is filled correctly (aspect 1).

$A_2 \equiv$ the customer is satisfied that the billing amount on the invoice is correct (aspect 2).

$\vdots$

$A_{10} \equiv$ the customer is satisfied that the delivery is made on time (aspect 10).

Also, define the event

$S \equiv$ the customer is satisfied with all 10 aspects of customer service.

Because 10 different departments are responsible for the 10 aspects of service being studied, it is reasonable to assume that all 10 aspects of service are independent of each other. For instance, billing amounts would be independent of delivery times. Therefore, $A_1, A_2, \ldots, A_{10}$ are independent events, and

$$\begin{aligned} P(S) &= P(A_1 \cap A_2 \cap \cdots \cap A_{10}) \\ &= P(A_1)P(A_2) \cdots P(A_{10}) \end{aligned}$$

If, as the departments claim, each department satisfies its customers 95 percent of the time, then the probability that the customer is satisfied with all 10 aspects is

$$P(S) = (.95)(.95) \cdots (.95) = (.95)^{10} = .5987$$

This result is almost identical to the 59 percent satisfaction rate reported by the survey participants.

If the company wants to increase the percentage of its customers who are satisfied with all 10 aspects of service, it must improve the quality of service provided by the 10 departments. For example, to satisfy 95 percent of its customers with all 10 aspects of service, the company must require each department to raise the fraction of the time it satisfies its customers to x, where x is such that $(x)^{10} = .95$. It follows that

$$x = (.95)^{\frac{1}{10}} = .9949$$

and that each department must satisfy its customers 99.49 percent of the time (rather than the current 95 percent of the time).

Exercises for Section 4.4

CONCEPTS

4.16 Give an example of a conditional probability that would be of interest to you.

4.17 Explain what it means for two events to be independent.

METHODS AND APPLICATIONS

4.18 The following contingency table summarizes the number of students at a college who have a Mastercard and/or a Visa credit card.

	Have Visa	Do Not Have Visa	Total
Have Mastercard	1,000	1,500	2,500
Do Not Have Mastercard	3,000	4,500	7,500
Total	4,000	6,000	10,000

a Find the proportion of Mastercard holders who have VISA cards. Interpret and write this proportion as a conditional probability.

b Find the proportion of VISA cardholders who have Mastercards. Interpret and write this proportion as a conditional probability.

c Are the events *having a Mastercard* and *having a VISA* independent? Justify your answer.

4.19 Each month a brokerage house studies various companies and rates each company's stock as being either "low risk" or "moderate to high risk." In a recent report, the brokerage house summarized its findings about 15 aerospace companies and 25 food retailers in the following table:

Company Type	Low Risk	Moderate to High Risk
Aerospace company	6	9
Food retailer	15	10

If we randomly select one of the total of 40 companies, find

a The probability that the company's stock is moderate to high risk given that the firm is an aerospace company.

b The probability that the company's stock is moderate to high risk given that the firm is a food retailer.

c Determine if the events *the firm is a food retailer* and *the firm's stock is low risk* are independent. Explain.

4.20 John and Jane are married. The probability that John watches a certain television show is .4. The probability that Jane watches the show is .5. The probability that John watches the show, given that Jane does, is .7.

a Find the probability that both John and Jane watch the show.

b Find the probability that Jane watches the show, given that John does.

c Do John and Jane watch the show independently of each other? Justify your answer.

4.21 In Exercise 4.20, find the probability that either John or Jane watches the show.

4.22 In the July 29, 2001 issue of the *Journal News* (Hamilton, Ohio), Lynn Elber of the Associated Press reported that "while 40 percent of American families own a television set with a V-chip installed to block designated programs with sex and violence, only 17 percent of those parents use the device."[4]

a Use the report's results to find an estimate of the probability that a randomly selected American family has used a V-chip to block programs containing sex and violence.

b According to the report, more than 50 percent of parents have used the TV rating system (TV-14, etc.) to control their children's TV viewing. How does this compare to the percentage using the V-chip?

4.23 According to the Associated Press report (in Exercise 4.22), 47 percent of parents who have purchased TV sets after V-chips became standard equipment in January 2000 are aware that their sets have V-chips, and of those who are aware of the option, 36 percent have programmed their V-chips. Using these results, find an estimate of the probability that a randomly selected parent who has bought a TV set since January 2000 has programmed the V-chip.

4.24 Fifteen percent of the employees in a company have managerial positions, and 25 percent of the employees in the company have MBA degrees. Also, 60 percent of the managers have MBA degrees. Using the probability formulas,

a Find the proportion of employees who are managers and have MBA degrees.

b Find the proportion of MBAs who are managers.

c Are the events *being a manager* and *having an MBA* independent? Justify your answer.

4.25 In Exercise 4.24, find the proportion of employees who either have MBAs or are managers.

4.26 Consider Exercise 4.14 (page 167). Using the results in Table 4.2 (page 167), estimate the probability that a randomly selected 21- to 49-year-old consumer would

a Give the phrase a rating of 4 or 5 given that the consumer is male; give the phrase a rating of 4 or 5 given that the consumer is female. Based on these results, is the appeal of the phrase among males much different from the appeal of the phrase among females? Explain.

b Give the phrase a rating of 4 or 5, given that the consumer is in the 21–24 age group; given that the consumer is in the 25–34 age group; given that the consumer is in the 35–49 age group. Based on these results, which age group finds the phrase most appealing? Least appealing?

4.27 In a survey of 100 insurance claims, 40 are fire claims (*FIRE*), 16 of which are fraudulent (*FRAUD*). Also, there are a total of 40 fraudulent claims.

a Construct a contingency table summarizing the claims data. Use the pairs of events *FIRE* and $\overline{FIRE}$, *FRAUD* and $\overline{FRAUD}$.

b What proportion of the fire claims are fraudulent?

c Are the events *a claim is fraudulent* and *a claim is a fire claim* independent? Use your probability of part *b* to prove your answer.

4.28 Recall from Exercise 4.3 (page 160) that two randomly selected customers are each asked to take a blind taste test and then to state which of three diet colas (marked as *A*, *B*, or *C*) he or she prefers. Suppose that cola *A*'s distributor claims that 80 percent of all people prefer cola *A* and that only 10 percent prefer each of colas *B* and *C*.

a Assuming that the distributor's claim is true and that the two taste test participants make independent cola preference decisions, find the probability of each sample space outcome.

b Find the probability that neither taste test participant will prefer cola *A*.

c If, when the taste test is carried out, neither participant prefers cola *A*, use the probability you computed in part *b* to decide whether the distributor's claim seems valid. Explain.

[4]Source: *Journal News* (Hamilton, Ohio), July 29, 2001, p. C5.

4.29 A sprinkler system inside an office building has two types of activation devices, $D1$ and $D2$, which operate independently. When there is a fire, if either device operates correctly, the sprinkler system is turned on. In case of fire, the probability that $D1$ operates correctly is .95, and the probability that $D2$ operates correctly is .92. Find the probability that

- **a** Both $D1$ and $D2$ will operate correctly.
- **b** The sprinkler system will come on.
- **c** The sprinkler system will fail.

4.30 A product is assembled using 10 different components, each of which must meet specifications for five different quality characteristics. Suppose that there is a .9973 probability that each individual specification will be met. Assuming that all 50 specifications are met independently, find the probability that the product meets all 50 specifications.

4.31 In Exercise 4.30, suppose that we wish to have a 99.73 percent chance that all 50 specifications will be met. If each specification will have the same chance of being met, how large must we make the probability of meeting each individual specification?

4.32 GENDER ISSUES AT A DISCOUNT CHAIN

Suppose that 65 percent of a discount chain's employees are women and 33 percent of the discount chain's employees having a management position are women. If 25 percent of the discount chain's employees have a management position, what percentage of the discount chain's female employees have a management position?

4.33 In a murder trial in Los Angeles, the prosecution claims that the defendant was cut on the left middle finger at the murder scene, but the defendant claims the cut occurred in Chicago, the day after the murders had been committed. Because the defendant is a sports celebrity, many people noticed him before he reached Chicago. Twenty-two people saw him casually, one person on the plane to Chicago carefully studied his hands looking for a championship ring, and another person stood with him as he signed autographs and drove him from the airport to the hotel. None of these 24 people saw a cut on the defendant's finger. If in fact he was not cut at all, it would be extremely unlikely that he left blood at the murder scene.

- **a** Because a person casually meeting the defendant would not be looking for a cut, assume that the probability is .9 that such a person would not have seen the cut, even if it was there. Furthermore, assume that the person who carefully looked at the defendant's hands had a .5 probability of not seeing the cut even if it was there and that the person who drove the defendant from the airport to the hotel had a .6 probability of not seeing the cut even if it was there. Given these assumptions, and also assuming that all 24 people looked at the defendant independently of each other, what is the probability that none of the 24 people would have seen the cut, even if it was there?
- **b** What is the probability that at least one of the 24 people would have seen the cut if it was there?
- **c** Given the result of part *b* and given the fact that none of the 24 people saw a cut, do you think the defendant had a cut on his hand before he reached Chicago?
- **d** How might we estimate what the assumed probabilities in *a* would actually be? (Note: This would not be easy.)

4.5 Bayes' Theorem (Optional)

LO4-5 Use Bayes' Theorem to update prior probabilities to posterior probabilities (Optional).

Sometimes we have an initial or **prior probability** that an event will occur. Then, based on new information, we revise the prior probability to what is called a **posterior probability.** This revision can be done by using a theorem called **Bayes' theorem.**

EXAMPLE 4.13 Should HIV Testing Be Mandatory?

HIV (Human Immunodeficiency Virus) is the virus that causes AIDS. Although many have proposed mandatory testing for HIV, statisticians have frequently spoken against such proposals. In this example, we use Bayes' theorem to see why.

Let HIV represent the event that a randomly selected American has the HIV virus, and let $\overline{HIV}$ represent the event that a randomly selected American does not have this virus. Because it is estimated that .6 percent of the American population have the HIV virus, $P(HIV) = .006$ and $P(\overline{HIV}) = .994$. A diagnostic test is used to attempt to detect whether a person has HIV. According to historical data, 99.9 percent of people with HIV receive a positive (POS) result when this test is administered, while 1 percent of people who do not have HIV receive a positive result. That is, $P(POS \mid HIV) = .999$ and $P(POS \mid \overline{HIV}) = .01$. If we administer the test to a randomly selected American (who may or may not have HIV) and the person receives a positive test result, what is the probability that the person actually has HIV? This probability is

$$P(HIV \mid POS) = \frac{P(HIV \cap POS)}{P(POS)}$$

The idea behind Bayes' theorem is that we can find $P(HIV \mid POS)$ by thinking as follows. A person will receive a positive result (POS) if the person receives a positive result and actually has HIV—that is, $(HIV \cap POS)$—or if the person receives a positive result and actually does not have HIV—that is, $(\overline{HIV} \cap POS)$. Therefore,

$$P(POS) = P(HIV \cap POS) + P(\overline{HIV} \cap POS)$$

This implies that

$$\begin{aligned} P(HIV \mid POS) &= \frac{P(HIV \cap POS)}{P(POS)} \\ &= \frac{P(HIV \cap POS)}{P(HIV \cap POS) + P(\overline{HIV} \cap POS)} \\ &= \frac{P(HIV)P(POS \mid HIV)}{P(HIV)P(POS \mid HIV) + P(\overline{HIV})P(POS \mid \overline{HIV})} \\ &= \frac{.006(.999)}{.006(.999) + (.994)(.01)} = .38 \end{aligned}$$

This probability says that, if all Americans were given a test for HIV, only 38 percent of the people who get a positive result would actually have HIV. That is, 62 percent of Americans identified as having HIV would actually be free of the virus! The reason for this rather surprising result is that, because so few people actually have HIV, the majority of people who test positive are people who are free of HIV and, therefore, erroneously test positive. This is why statisticians have spoken against proposals for mandatory HIV testing.

In the preceding example, there were two *states of nature*—HIV and $\overline{HIV}$—and two outcomes of the diagnostic test—POS and $\overline{POS}$. In general, there might be any number of states of nature and any number of experimental outcomes. This leads to a general statement of Bayes' theorem.

Bayes' Theorem

Let $S_1, S_2, \ldots, S_k$ be k mutually exclusive states of nature, one of which must be true, and suppose that $P(S_1), P(S_2), \ldots, P(S_k)$ are the prior probabilities of these states of nature. Also, let E be a particular outcome of an experiment designed to help determine which state of nature is really true. Then, the **posterior probability** of a particular state of nature, say S_i, given the experimental outcome E, is

$$P(S_i \mid E) = \frac{P(S_i \cap E)}{P(E)} = \frac{P(S_i)P(E \mid S_i)}{P(E)}$$

where

$$\begin{aligned} P(E) &= P(S_1 \cap E) + P(S_2 \cap E) + \cdots + P(S_k \cap E) \\ &= P(S_1)P(E \mid S_1) + P(S_2)P(E \mid S_2) + \cdots + P(S_k)P(E \mid S_k) \end{aligned}$$

Specifically, if there are two mutually exclusive states of nature, S_1 and S_2, one of which must be true, then

$$P(S_i \mid E) = \frac{P(S_i)P(E \mid S_i)}{P(S_1)P(E \mid S_1) + P(S_2)P(E \mid S_2)}$$

We have illustrated Bayes' theorem when there are two states of nature in Example 4.13. In the next example, we consider three states of nature.

EXAMPLE 4.14 The Oil Drilling Case: Site Selection

C

An oil company is attempting to decide whether to drill for oil on a particular site. There are three possible states of nature:

1 No oil (state of nature S_1, which we will denote as *none*)
2 Some oil (state of nature S_2, which we will denote as *some*)
3 Much oil (state of nature S_3, which we will denote as *much*)

Based on experience and knowledge concerning the site's geological characteristics, the oil company feels that the prior probabilities of these states of nature are as follows:

$$P(S_1 \equiv \text{none}) = .7 \qquad P(S_2 \equiv \text{some}) = .2 \qquad P(S_3 \equiv \text{much}) = .1$$

In order to obtain more information about the potential drilling site, the oil company can perform a seismic experiment, which has three readings—low, medium, and high. Moreover, information exists concerning the accuracy of the seismic experiment. The company's historical records tell us that

1 Of 100 past sites that were drilled and produced no oil, 4 sites gave a high reading. Therefore,

$$P(\text{high} \mid \text{none}) = \frac{4}{100} = .04$$

2 Of 400 past sites that were drilled and produced some oil, 8 sites gave a high reading. Therefore,

$$P(\text{high} \mid \text{some}) = \frac{8}{400} = .02$$

3 Of 300 past sites that were drilled and produced much oil, 288 sites gave a high reading. Therefore,

$$P(\text{high} \mid \text{much}) = \frac{288}{300} = .96$$

Intuitively, these conditional probabilities tell us that sites that produce no oil or some oil seldom give a high reading, while sites that produce much oil often give a high reading.

Now, suppose that when the company performs the seismic experiment on the site in question, it obtains a high reading. The previously given conditional probabilities suggest that, given this new information, the company might feel that the likelihood of much oil is higher than its prior probability $P(\text{much}) = .1$, and that the likelihoods of some oil and no oil are lower than the prior probabilities $P(\text{some}) = .2$ and $P(\text{none}) = .7$. To be more specific, we wish to *revise the prior probabilities* of no, some, and much oil to what we call *posterior probabilities.* We can do this by using Bayes' theorem as follows.

If we wish to compute $P(\text{none} \mid \text{high})$, we first calculate

$$\begin{aligned} P(\text{high}) &= P(\text{none} \cap \text{high}) + P(\text{some} \cap \text{high}) + P(\text{much} \cap \text{high}) \\ &= P(\text{none})P(\text{high} \mid \text{none}) + P(\text{some})P(\text{high} \mid \text{some}) + P(\text{much})P(\text{high} \mid \text{much}) \\ &= (.7)(.04) + (.2)(.02) + (.1)(.96) = .128 \end{aligned}$$

Then Bayes' theorem says that

$$P(\text{none} \mid \text{high}) = \frac{P(\text{none} \cap \text{high})}{P(\text{high})} = \frac{P(\text{none})P(\text{high} \mid \text{none})}{P(\text{high})} = \frac{.7(.04)}{.128} = .21875$$

Similarly, we can compute $P(\text{some} \mid \text{high})$ and $P(\text{much} \mid \text{high})$ as follows.

$$P(\text{some} \mid \text{high}) = \frac{P(\text{some} \cap \text{high})}{P(\text{high})} = \frac{P(\text{some})P(\text{high} \mid \text{some})}{P(\text{high})} = \frac{.2(.02)}{.128} = .03125$$

$$P(\text{much} \mid \text{high}) = \frac{P(\text{much} \cap \text{high})}{P(\text{high})} = \frac{P(\text{much})P(\text{high} \mid \text{much})}{P(\text{high})} = \frac{.1(.96)}{.128} = .75$$

These revised probabilities tell us that, given that the seismic experiment gives a high reading, the revised probabilities of no, some, and much oil are .21875, .03125, and .75, respectively.

Because the posterior probability of much oil is .75, we might conclude that we should drill on the oil site. However, this decision should also be based on economic considerations. The science of **decision theory** provides various criteria for making such a decision. An introduction to decision theory can be found in Chapter 19.

In this section we have only introduced Bayes' theorem. There is an entire subject called **Bayesian statistics,** which uses Bayes' theorem to update prior belief about a probability or population parameter to posterior belief. The use of Bayesian statistics is controversial in the case where the prior belief is largely based on subjective considerations, because many statisticians do not believe that we should base decisions on subjective considerations. Realistically, however, we all do this in our daily lives. For example, how each of us viewed the evidence in the O. J. Simpson murder trial had a great deal to do with our prior beliefs about both O. J. Simpson and the police.

Exercises for Section 4.5

CONCEPTS

4.34 What is a prior probability? What is a posterior probability?

4.35 Explain the purpose behind using Bayes' theorem.

METHODS AND APPLICATIONS

4.36 Suppose that A_1, A_2, and B are events where A_1 and A_2 are mutually exclusive and

$$P(A_1) = .8 \qquad P(B \mid A_1) = .1$$
$$P(A_2) = .2 \qquad P(B \mid A_2) = .3$$

Use this information to find $P(A_1 \mid B)$ and $P(A_2 \mid B)$.

4.37 Suppose that A_1, A_2, A_3, and B are events where A_1, A_2, and A_3 are mutually exclusive and

$$P(A_1) = .2 \qquad P(A_2) = .5 \qquad P(A_3) = .3$$
$$P(B \mid A_1) = .02 \qquad P(B \mid A_2) = .05 \qquad P(B \mid A_3) = .04$$

Use this information to find $P(A_1 \mid B)$, $P(A_2 \mid B)$ and $P(A_3 \mid B)$.

4.38 Again consider the diagnostic test for HIV discussed in Example 4.13 (page 175) and recall that $P(POS \mid HIV) = .999$ and $P(POS \mid \overline{HIV}) = .01$, where POS denotes a positive test result. Assuming that the percentage of people who have HIV is 1 percent, recalculate the probability that a randomly selected person has HIV, given that his or her test result is positive.

4.39 A department store is considering a new credit policy to try to reduce the number of customers defaulting on payments. A suggestion is made to discontinue credit to any customer who has been one week or more late with his/her payment at least twice. Past records show 95 percent of defaults were late at least twice. Also, 3 percent of all customers default, and 30 percent of those who have not defaulted have had at least two late payments.

a Find the probability that a customer with at least two late payments will default.
b Based on part *a*, should the policy be adopted? Explain.

4.40 A company administers an "aptitude test for managers" to aid in selecting new management trainees. Prior experience suggests that 60 percent of all applicants for management trainee positions would be successful if they were hired. Furthermore, past experience with the aptitude test indicates that 85 percent of applicants who turn out to be successful managers pass the test and 90 percent of applicants who do not turn out to be successful managers fail the test.

a If an applicant passes the "aptitude test for managers," what is the probability that the applicant will succeed in a management position?
b Based on your answer to part *a*, do you think that the "aptitude test for managers" is a valuable way to screen applicants for management trainee positions? Explain.

4.41 THE OIL DRILLING CASE

Recall that the prior probabilities of no oil (*none*), some oil (*some*), and much oil (*much*) are:

$$P(\text{none}) = .7 \qquad P(\text{some}) = .2 \qquad P(\text{much}) = .1$$

Of 100 past sites that were drilled and produced no oil, 5 gave a medium reading. Of the 400 past sites that were drilled and produced some oil, 376 gave a medium reading. Of the 300 past sites that were drilled and produced much oil, 9 gave a medium reading. This implies that the conditional probabilities of a medium reading (medium) given no oil, some oil, and much oil are:

$$P(\text{medium} \mid \text{none}) = \frac{5}{100} = .05$$

$$P(\text{medium} \mid \text{some}) = \frac{376}{400} = .94$$

$$P(\text{medium} \mid \text{much}) = \frac{9}{300} = .03$$

Calculate the posterior probabilities of no, some, and much oil, given a medium reading.

4.42 THE OIL DRILLING CASE

Of 100 past sites that were drilled and produced no oil, 91 gave a low reading. Of the 400 past sites that were drilled and produced some oil, 16 gave a low reading. Of the 300 past sites that were drilled and produced much oil, 3 gave a low reading. Calculate the posterior probabilities of no, some, and much oil, given a low reading.

4.43 Three data entry specialists enter requisitions into a computer. Specialist 1 processes 30 percent of the requisitions, specialist 2 processes 45 percent, and specialist 3 processes 25 percent. The proportions of incorrectly entered requisitions by data entry specialists 1, 2, and 3 are .03, .05, and .02, respectively. Suppose that a random requisition is found to have been incorrectly entered. What is the probability that it was processed by data entry specialist 1? By data entry specialist 2? By data entry specialist 3?

4.44 A truth serum given to a suspect is known to be 90 percent reliable when the person is guilty and 99 percent reliable when the person is innocent. In other words, 10 percent of the guilty are judged innocent by the serum and 1 percent of the innocent are judged guilty. If the suspect was selected from a group of suspects of which only 5 percent are guilty of having committed a crime, and the serum indicates that the suspect is guilty of having committed a crime, what is the probability that the suspect is innocent?

4.6 Counting Rules (Optional)

LO4-6 Use elementary counting rules to compute probabilities (Optional).

Consider the situation in Example 4.2 (page 156) in which a student takes a pop quiz that consists of three true–false questions. If we consider our experiment to be answering the three questions, each question can be answered correctly or incorrectly. We will let C denote answering a question correctly and I denote answering a question incorrectly. Figure 4.6 depicts a tree diagram of the sample space outcomes for the experiment. The diagram portrays the experiment as a three-step process—answering the first question (correctly or incorrectly, that is, C or I), answering the second question (correctly or incorrectly, that is, C or I), and answering the third question (correctly or incorrectly, that is, C or I). The tree diagram has eight different branches, and the eight distinct sample space outcomes are listed at the ends of the branches.

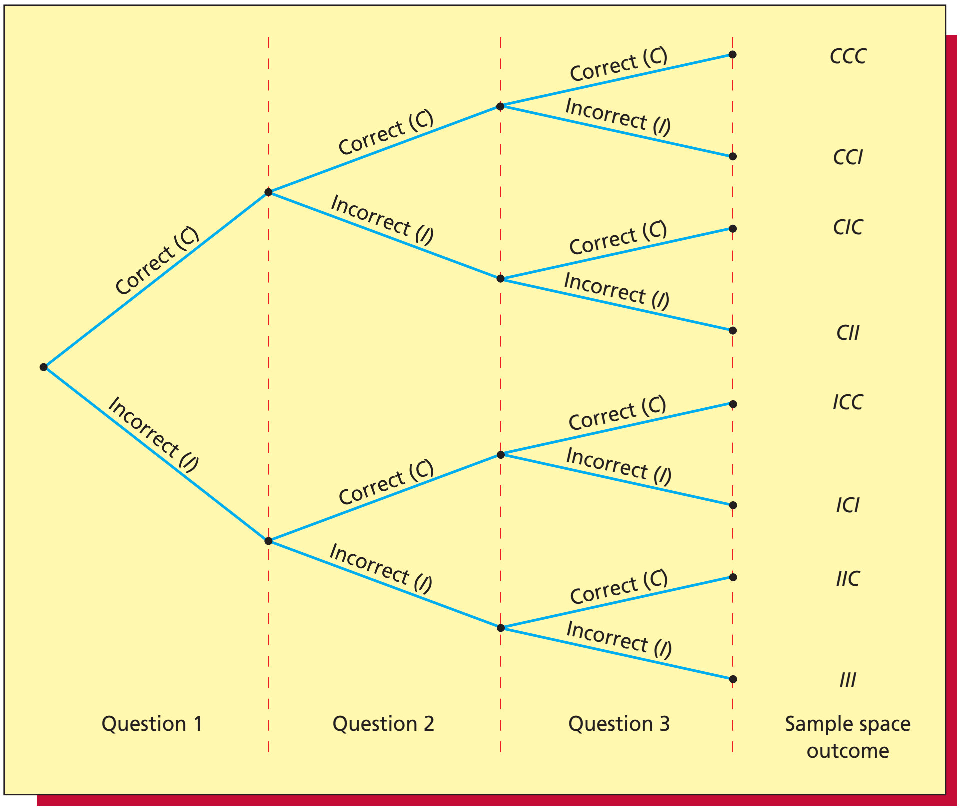
In general, a rule that is helpful in determining the number of experimental outcomes in a multiple-step experiment is as follows:

A Counting Rule for Multiple-Step Experiments

If an experiment can be described as a sequence of k steps in which there are n_1 possible outcomes on the first step, n_2 possible outcomes on the second step, and so on, then the total number of experimental outcomes is given by $(n_1)(n_2) \cdots (n_k)$.

For example, the pop quiz example consists of three steps in which there are $n_1 = 2$ possible outcomes on the first step, $n_2 = 2$ possible outcomes on the second step, and $n_3 = 2$ possible outcomes on the third step. Therefore, the total number of experimental outcomes is $(n_1)(n_2)(n_3) = (2)(2)(2) = 8$, as is shown in Figure 4.6. Now suppose the student takes a pop quiz consisting of

FIGURE 4.6 A Tree Diagram of Answering Three True–False Questions



five true–false questions. Then, there are $(n_1)(n_2)(n_3)(n_4)(n_5) = (2)(2)(2)(2)(2) = 32$ experimental outcomes. If the student is totally unprepared for the quiz and has to blindly guess the answer to each question, the 32 experimental outcomes might be considered to be equally likely. Therefore, since only one of these outcomes corresponds to all five questions being answered correctly, the probability that the student will answer all five questions correctly is $1/32$.

As another example, suppose a bank has three branches; each branch has two departments, and each department has four employees. One employee is to be randomly selected to go to a convention. Because there are $(n_1)(n_2)(n_3) = (3)(2)(4) = 24$ employees, the probability that a particular one will be randomly selected is $1/24$.

Next, consider the population of last year's percentage returns for six high-risk stocks. This population consists of the percentage returns -36, -15, 3, 15, 33, and 54 (which we have arranged in increasing order). Now consider randomly selecting without replacement a sample of $n = 3$ stock returns from the population of six stock returns. Below we list the 20 distinct samples of $n = 3$ returns that can be obtained:

Sample	$n = 3$ Returns in Sample	Sample	$n = 3$ Returns in Sample
1	−36, −15, 3	11	−15, 3, 15
2	−36, −15, 15	12	−15, 3, 33
3	−36, −15, 33	13	−15, 3, 54
4	−36, −15, 54	14	−15, 15, 33
5	−36, 3, 15	15	−15, 15, 54
6	−36, 3, 33	16	−15, 33, 54
7	−36, 3, 54	17	3, 15, 33
8	−36, 15, 33	18	3, 15, 54
9	−36, 15, 54	19	3, 33, 54
10	−36, 33, 54	20	15, 33, 54

Because each sample is specified only with respect to which returns are contained in the sample, and therefore not with respect to the different orders in which the returns can be randomly selected, each sample is called a **combination of $n = 3$ stock returns selected from $N = 6$ stock returns.** In general, the following result can be proven:

A Counting Rule for Combinations

The number of combinations of n items that can be selected from N items is

$$\binom{N}{n} = \frac{N!}{n!\,(N-n)!}$$

where

$$N! = N(N-1)(N-2)\cdots 1$$

$$n! = n(n-1)(n-2)\cdots 1$$

Note: 0! is defined to be 1.

For example, the number of combinations of $n = 3$ stock returns that can be selected from the six previously discussed stock returns is

$$\binom{6}{3} = \frac{6!}{3!(6-3)!} = \frac{6!}{3!\,3!} = \frac{6\cdot 5\cdot 4\,(\cancel{3\cdot 2\cdot 1})}{(3\cdot 2\cdot 1)\,(\cancel{3\cdot 2\cdot 1})} = 20$$

The 20 combinations are listed on the previous page. As another example, the Ohio lottery system uses the random selection of 6 numbers from a group of 47 numbers to determine each week's lottery winner. There are

$$\binom{47}{6} = \frac{47!}{6!\,(47-6)!} = \frac{47\cdot 46\cdot 45\cdot 44\cdot 43\cdot 42\,(\cancel{41!})}{6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1\,(\cancel{41!})} = 10{,}737{,}573$$

combinations of 6 numbers that can be selected from 47 numbers. Therefore, if you buy a lottery ticket and pick six numbers, the probability that this ticket will win the lottery is 1/10,737,573.

Exercises for Section 4.6

CONCEPTS

4.45 Explain why counting rules are useful.

4.46 Explain when it is appropriate to use the counting rule for multiple-step experiments.

4.47 Explain when it is appropriate to use the counting rule for combinations.

METHODS AND APPLICATIONS

4.48 A credit union has two branches; each branch has two departments, and each department has four employees. How many total people does the credit union employ? If you work for the credit union, and one employee is randomly selected to go to a convention, what is the probability that you will be chosen?

4.49 Construct a tree diagram (like Figure 4.6) for the situation described in Exercise 4.48.

4.50 How many combinations of two high-risk stocks could you randomly select from eight high-risk stocks? If you did this, what is the probability that you would obtain the two highest-returning stocks?

4.51 A pop quiz consists of three true–false questions and three multiple choice questions. Each multiple choice question has five possible answers. If a student blindly guesses the answer to every question, what is the probability that the student will correctly answer all six questions?

4.52 A company employs eight people and plans to select a group of three of these employees to receive advanced training. How many ways can the group of three employees be selected?

4.53 The company of Exercise 4.52 employs Mr. Withrow, Mr. Church, Ms. David, Ms. Henry, Mr. Fielding, Mr. Smithson, Ms. Penny, and Mr. Butler. If the three employees who will receive advanced training are selected at random, what is the probability that Mr. Church, Ms. Henry, and Mr. Butler will be selected for advanced training?

Chapter Summary

In this chapter we studied **probability.** We began by defining an **event** to be an experimental outcome that may or may not occur and by defining the **probability of an event** to be a number that measures the likelihood that the event will occur. We learned that a probability is often interpreted as a **long-run relative frequency,** and we saw that probabilities can be found by examining **sample spaces** and by using **probability rules.** We learned several important probability rules—**addition rules, multiplication rules,** and **the rule of complements.** We also studied a special kind of probability called a **conditional probability,** which is the probability that one event will occur given that another event occurs, and we used probabilities to define **independent events.** We concluded this chapter by studying two optional topics. The first of these was **Bayes' theorem,** which can be used to update a **prior** probability to a **posterior** probability based on receiving new information. Second, we studied **counting rules** that are helpful when we wish to count sample space outcomes.

Glossary of Terms

Bayes' theorem: A theorem (formula) that is used to compute posterior probabilities by revising prior probabilities. (page 176)
Bayesian statistics: An area of statistics that uses Bayes' Theorem to update prior belief about a probability or population parameter to posterior belief. (page 178)
complement (of an event): If A is an event, the complement of A is the event that A will not occur. (page 161)
conditional probability: The probability that one event will occur given that we know that another event occurs. (page 168)
decision theory: An approach that helps decision makers to make intelligent choices. (page 178)
dependent events: When the probability of one event is influenced by whether another event occurs, the events are said to be dependent. (page 171)
event: A set of one or more sample space outcomes. (page 155)
experiment: A process of observation that has an uncertain outcome. (page 153)
independent events: When the probability of one event is not influenced by whether another event occurs, the events are said to be independent. (page 171)
mutually exclusive events: Events that have no sample space outcomes in common, and, therefore, cannot occur simultaneously. (page 164)
prior probability: The initial probability that an event will occur. (page 175)
probability (of an event): A number that measures the chance, or likelihood, that an event will occur when an experiment is carried out. (pages 153, 155)
posterior probability: A revised probability obtained by updating a prior probability after receiving new information. (page 175)
sample space: The set of all possible experimental outcomes (sample space outcomes). (page 153)
sample space outcome: A distinct outcome of an experiment (that is, an element in the sample space). (page 153)
subjective probability: A probability assessment that is based on experience, intuitive judgment, or expertise. (page 155)

Important Formulas

Probabilities when all sample space outcomes are equally likely: page 158

The rule of complements: page 161

The addition rule for two events: page 164

The addition rule for two mutually exclusive events: page 164

The addition rule for N mutually exclusive events: page 166

Conditional probability: pages 168, 169

The general multiplication rule: page 169

Independence: page 171

The multiplication rule for two independent events: page 172

The multiplication rule for N independent events: page 172

Bayes' theorem: page 176

Counting rule for multiple-step experiments: page 179

Counting rule for combinations: page 181

Supplementary Exercises

Exercises 4.54 through 4.57 are based on the following situation: An investor holds two stocks, each of which can rise (R), remain unchanged (U), or decline (D) on any particular day.

connect

4.54 Construct a tree diagram showing all possible combined movements for both stocks on a particular day (for instance, RR, RD, and so on, where the first letter denotes the movement of the first stock, and the second letter denotes the movement of the second stock).

4.55 If all outcomes are equally likely, find the probability that both stocks rise; that both stocks decline; that exactly one stock declines.

4.56 Find the probabilities you found in Exercise 4.55 by assuming that for each stock $P(R) = .6$, $P(U) = .1$, and $P(D) = .3$, and assuming that the two stocks move independently.

4.57 Assume that for the first stock (on a particular day)

$$P(R) = .4,\ P(U) = .2,\ P(D) = .4$$

and that for the second stock (on a particular day)

$$P(R) = .8,\ P(U) = .1,\ P(D) = .1$$

Assuming that these stocks move independently, find the probability that both stocks decline; the probability that exactly one stock rises; the probability that exactly one stock is unchanged; the probability that both stocks rise.

The Bureau of Labor Statistics reports on a variety of employment statistics. "College Enrollment and Work Activity of 2004 High School Graduates" provides information on high school graduates by gender, by race, and by labor force participation as of October 2004.[5] (All numbers are in thousands.) The following two tables provide sample information on the "Labor force status of persons 16 to 24 years old by educational attainment and gender, October 2004." Using the information contained in the tables, do Exercises 4.58 through 4.62. DS LabForce

Women, Age 16 to 24	Civilian Labor Force Employed	Civilian Labor Force Unemployed	Not in Labor Force	Row Total
< High School	662	205	759	1,626
HS degree	2,050	334	881	3,265
Some college	1,352	126	321	1,799
Bachelors degree or more	921	55	105	1,081
Column Total	4,985	720	2,066	7,771

Men, Age 16 to 24	Civilian Labor Force Employed	Civilian Labor Force Unemployed	Not in Labor Force	Row Total
< High School	1,334	334	472	2,140
HS degree	3,110	429	438	3,977
Some college	1,425	106	126	1,657
Bachelors degree or more	708	37	38	783
Column Total	6,577	906	1,074	8,557

4.58 Find an estimate of the probability that a randomly selected female aged 16 to 24 is in the civilian labor force, if she has a high school degree. DS LabForce

4.59 Find an estimate of the probability that a randomly selected female aged 16 to 24 is in the civilian labor force, if she has a bachelors degree or more. DS LabForce

4.60 Find an estimate of the probability that a randomly selected female aged 16 to 24 is employed, if she is in the civilian labor force and has a high school degree. DS LabForce

4.61 Find an estimate of the probability that a randomly selected female aged 16 to 24 is employed, if she is in the civilian labor force and has a bachelor's degree or more. DS LabForce

4.62 Repeat Exercises 4.58 through 4.61 for a randomly selected male aged 16 to 24. In general, do the tables imply that labor force status and employment status depend upon educational attainment? Explain your answer. DS LabForce

Suppose that in a survey of 1,000 U.S. residents, 721 residents believed that the amount of violent television programming had increased over the past 10 years, 454 residents believed that the overall quality of television programming had decreased over the past 10 years, and 362 residents believed both. Use this information to do Exercises 4.63 through 4.69.

4.63 What proportion of the 1,000 U.S. residents believed that the amount of violent programming had increased over the past 10 years?

[5] Source: *College Enrollment and Work Activity of 2004 High School Graduates,* Table 2, "Labor force status of persons 16 to 24 years old by school enrollment, educational attainment, sex, race, and Hispanic or Latino ethnicity, October 2004," www.bls.gov.

4.64 What proportion of the 1,000 U.S. residents believed that the overall quality of programming had decreased over the past 10 years?

4.65 What proportion of the 1,000 U.S. residents believed that both the amount of violent programming had increased and the overall quality of programming had decreased over the past 10 years?

4.66 What proportion of the 1,000 U.S. residents believed that either the amount of violent programming had increased or the overall quality of programming had decreased over the past 10 years?

4.67 What proportion of the U.S. residents who believed that the amount of violent programming had increased believed that the overall quality of programming had decreased?

4.68 What proportion of the U.S. residents who believed that the overall quality of programming had decreased believed that the amount of violent programming had increased?

4.69 What sort of dependence seems to exist between whether U.S. residents believed that the amount of violent programming had increased and whether U.S. residents believed that the overall quality of programming had decreased? Explain your answer.

4.70 Enterprise Industries has been running a television advertisement for Fresh liquid laundry detergent. When a survey was conducted, .21 of the individuals surveyed had purchased Fresh, .41 of the individuals surveyed had recalled seeing the advertisement, and .13 of the individuals surveyed had purchased Fresh and recalled seeing the advertisement.

a What proportion of the individuals surveyed who recalled seeing the advertisement had purchased Fresh?

b Based on your answer to part *a*, does the advertisement seem to have been effective? Explain.

4.71 A company employs 400 salespeople. Of these, 83 received a bonus last year, 100 attended a special sales training program at the beginning of last year, and 42 both attended the special sales training program and received a bonus. (Note: the bonus was based totally on sales performance.)

a What proportion of the 400 salespeople received a bonus last year?

b What proportion of the 400 salespeople attended the special sales training program at the beginning of last year?

c What proportion of the 400 salespeople both attended the special sales training program and received a bonus?

d What proportion of the salespeople who attended the special sales training program received a bonus?

e Based on your answers to parts *a* and *d*, does the special sales training program seem to have been effective? Explain your answer.

4.72 On any given day, the probability that the Ohio River at Cincinnati is polluted by a carbon tetrachloride spill is .10. Each day, a test is conducted to determine whether the river is polluted by carbon tetrachloride. This test has proved correct 80 percent of the time. Suppose that on a particular day the test indicates carbon tetrachloride pollution. What is the probability that such pollution actually exists?

4.73 In the book *Making Hard Decisions: An Introduction to Decision Analysis,* Robert T. Clemen presents an example in which he discusses the 1982 John Hinckley trial. In describing the case, Clemen says:

> In 1982 John Hinckley was on trial, accused of having attempted to kill President Reagan. During Hinckley's trial, Dr. Daniel R. Weinberger told the court that when individuals diagnosed as schizophrenics were given computerized axial tomography (CAT) scans, the scans showed brain atrophy in 30% of the cases compared with only 2% of the scans done on normal people. Hinckley's defense attorney wanted to introduce as evidence Hinckley's CAT scan, which showed brain atrophy. The defense argued that the presence of atrophy strengthened the case that Hinckley suffered from mental illness.

a Approximately 1.5 percent of the people in the United States suffer from schizophrenia. If we consider the prior probability of schizophrenia to be .015, use the information given to find the probability that a person has schizophrenia given that a person's CAT scan shows brain atrophy.

b John Hinckley's CAT scan showed brain atrophy. Discuss whether your answer to part *a* helps or hurts the case that Hinckley suffered from mental illness.

c It can be argued that .015 is not a reasonable prior probability of schizophrenia. This is because .015 is the probability that a randomly selected U.S. citizen has schizophrenia. However, John Hinckley was not a randomly selected U.S. citizen. Rather, he was accused of attempting to assassinate the president. Therefore, it might be reasonable to assess a higher prior probability of schizophrenia. Suppose you are a juror who believes there is only a 10 percent chance that Hinckley suffers from schizophrenia. Using .10 as the prior probability of schizophrenia, find the probability that a person has schizophrenia given that a person's CAT scan shows brain atrophy.

d If you are a juror with a prior probability of .10 that John Hinckley suffers from schizophrenia and given your answer to part *c*, does the fact that Hinckley's CAT scan showed brain atrophy help the case that Hinckley suffered from mental illness?

e If you are a juror with a prior probability of .25 that Hinckley suffers from schizophrenia, find the probability of schizophrenia given that Hinckley's CAT scan showed brain atrophy. In this situation, how strong is the case that Hinckley suffered from mental illness?

4.74 Below we give two contingency tables of data from reports submitted by airlines to the U.S. Department of Transportation. The data concern the numbers of on-time and delayed flights for Alaska Airlines and America West Airlines at five major airports. DS AirDelays

	Alaska Airlines				America West		
	On Time	Delayed	Total		On Time	Delayed	Total
Los Angeles	497	62	559	**Los Angeles**	694	117	811
Phoenix	221	12	233	**Phoenix**	4,840	415	5,255
San Diego	212	20	232	**San Diego**	383	65	448
San Francisco	503	102	605	**San Francisco**	320	129	449
Seattle	1,841	305	2,146	**Seattle**	201	61	262
Total	3,274	501	3,775	**Total**	6,438	787	7,225

Source: A. Barnett, "How Numbers Can Trick You," *Technology Review,* October 1994, pp. 38–45.

a What percentage of all Alaska Airlines flights were delayed? That is, use the data to estimate the probability that an Alaska Airlines flight will be delayed. Do the same for America West Airlines. Which airline does best overall?

b For Alaska Airlines, find the percentage of delayed flights at each airport. That is, use the data to estimate each of the probabilities $P(\text{delayed} \mid \text{Los Angeles})$, $P(\text{delayed} \mid \text{Phoenix})$, and so on. Then do the same for America West Airlines. Which airline does best at each individual airport?

c We find that America West Airlines does worse at every airport, yet America West does best overall. This seems impossible, but it is true! By looking carefully at the data, explain how this can happen. Hint: Consider the weather in Phoenix and Seattle. (This exercise is an example of what is called *Simpson's paradox.*)

4.75 Internet Exercise

What is the age, gender, and ethnic composition of U.S. college students? As background for its 1995 study of college students and their risk behaviors, the Centers for Disease Control and Prevention collected selected demographic data—age, gender, and ethnicity—about college students. A report on the 1995 National Health Risk Behavior Survey can be found at the CDC website by going directly to http://www.cdc.gov/mmwr/preview/mmwrhtml/00049859.htm. This report includes a large number of tables, the first of which summarizes the demographic information for the sample of $n = 4609$ college students. An excerpt from Table 1 is given on the right.

Using conditional probabilities, discuss (a) the dependence between age and gender and (b) the dependence between age and ethnicity for U.S. college students. DS CDCData

TABLE 1. Demographic Characteristics of Undergraduate College Students Aged >=18 Years, by Age Group — United States, National College Health Risk Behavior Survey, 1995

		Age Group (%)	
Category	Total (%)	18-24 Years	>=25 Years
Total	--	63.6	36.4
Sex			
Female	55.5	52.0	61.8
Male	44.5	48.0	38.2
Race/ethnicity			
White*	72.8	70.9	76.1
Black*	10.3	10.5	9.6
Hispanic	7.1	6.9	7.4
Other	9.9	11.7	6.9