Applying Equation (4.3) to the previous example produces the following result:

$$\begin{aligned}
P(\text{Planned to purchase } \textit{or} \text{ actually purchased}) &= P(\text{Planned to purchase}) \\
&\quad + P(\text{Actually purchased}) - P(\text{Planned to purchase } \textit{and} \text{ actually purchased}) \\
&= \frac{250}{1,000} + \frac{300}{1,000} - \frac{200}{1,000} \\
&= \frac{350}{1,000} = 0.35
\end{aligned}$$

The general addition rule consists of taking the probability of A and adding it to the probability of B and then subtracting the probability of the joint event A *and* B from this total because the joint event has already been included in computing both the probability of A and the probability of B. Referring to Table 4.1 on page 154, if the outcomes of the event "planned to purchase" are added to those of the event "actually purchased," the joint event "planned to purchase *and* actually purchased" has been included in each of these simple events. Therefore, because this joint event has been included twice, you must subtract it to compute the correct result. Example 4.5 illustrates another application of the general addition rule.

EXAMPLE 4.5

Using the General Addition Rule for the Households That Purchased Large-Screen HDTVs

In Example 4.3 on page 156, the purchases were cross-classified in Table 4.2 as televisions that had a faster refresh rate or televisions that had a standard refresh rate and whether the household purchased a streaming media box. Find the probability that among households that purchased a large-screen HDTV, they purchased a television that had a faster refresh rate or purchased a streaming media box.

SOLUTION Using Equation (4.3),

$$\begin{aligned}
&P(\text{Television had a faster refresh rate } \textit{or} \text{ purchased a streaming media box}) \\
&\quad = P(\text{Television had a faster refresh rate}) + P(\text{purchased a streaming media box}) \\
&\quad\quad - P(\text{Television had a faster refresh rate } \textit{and} \text{ purchased a streaming media box}) \\
&\quad = \frac{80}{300} + \frac{108}{300} - \frac{38}{300} \\
&\quad = \frac{150}{300} = 0.50
\end{aligned}$$

Therefore, of households that purchased a large-screen HDTV, there is a 50% chance that a randomly selected household purchased a television that had a faster refresh rate or purchased a streaming media box.

Problems for Section 4.1

LEARNING THE BASICS

4.1 Two coins are tossed.

a. Give an example of a simple event.
b. Give an example of a joint event.
c. What is the complement of a head on the first toss?
d. What does the sample space consist of?

4.2 An urn contains 12 red balls and 8 white balls. One ball is to be selected from the urn.

a. Give an example of a simple event.
b. What is the complement of a red ball?
c. What does the sample space consist of?

4.3 Consider the following contingency table:

	B	B'
A	10	20
A'	20	40

What is the probability of event
a. A?
b. A'?
c. A *and* B?
d. A *or* B?

4.4 Consider the following contingency table:

	B	B'
A	10	30
A'	25	35

What is the probability of event
a. A'?
b. A *and* B?
c. A' *and* B'?
d. A' *or* B'?

APPLYING THE CONCEPTS

4.5 For each of the following, indicate whether the type of probability involved is an example of *a priori* probability, empirical probability, or subjective probability.
a. The next toss of a fair coin will land on heads.
b. Italy will win soccer's World Cup the next time the competition is held.
c. The sum of the faces of two dice will be seven.
d. The train taking a commuter to work will be more than 10 minutes late.

4.6 For each of the following, state whether the events created are mutually exclusive and whether they are collectively exhaustive.
a. Undergraduate business students were asked whether they were sophomores or juniors.
b. Each respondent was classified by the type of car he or she drives: sedan, SUV, American, European, Asian, or none.
c. People were asked, "Do you currently live in (i) an apartment or (ii) a house?"
d. A product was classified as defective or not defective.

4.7 Which of the following events occur with a probability of zero? For each, state why or why not.
a. A company is listed on the New York Stock Exchange and NASDAQ.
b. A consumer owns a smartphone and a tablet.
c. A cellphone is a Motorola and a Samsung.
d. An automobile is a Toyota and was manufactured in the United States.

4.8 Do males or females feel more tense or stressed out at work? A survey of employed adults conducted online by Harris Interactive on behalf of the American Psychological Association revealed the following:

	FELT TENSE OR STRESSED OUT AT WORK	
GENDER	**Yes**	**No**
Male	244	495
Female	282	480

Source: Data extracted from "The 2013 Work and Well-Being Survey," American Psychological Association and Harris Interactive, March 2013, p. 5, **bit.ly/11JGcPf**.

a. Give an example of a simple event.
b. Give an example of a joint event.
c. What is the complement of "Felt tense or stressed out at work"?
d. Why is "Male and felt tense or stressed out at work" a joint event?

4.9 Referring to the contingency table in Problem 4.8, if an employed adult is selected at random, what is the probability that
a. the employed adult felt tense or stressed out at work?
b. the employed adult was a male who felt tense or stressed out at work?
c. the employed adult was a male *or* felt tense or stressed out at work?
d. Explain the difference in the results in (b) and (c).

4.10 How will marketers change their social media use in the near future? A survey by Social Media Examiner reported that 76% of B2B marketers (marketers that focus primarily on attracting businesses) plan to increase their use of LinkedIn, as compared to 55% of B2C marketers (marketers that primarily target consumers). The survey was based on 1,945 B2B marketers and 1,868 B2C marketers. The following table summarizes the results:

	BUSINESS FOCUS		
INCREASE USE OF LINKEDIN?	**B2B**	**B2C**	**Total**
Yes	1,478	1,027	2,505
No	467	841	1,308
Total	1,945	1,868	3,813

Source: Data extracted from "2012 Social Media Marketing Industry Report," April 2012, p. 27, **bit.ly/HaWwDu**.

a. Give an example of a simple event.
b. Give an example of a joint event.
c. What is the complement of a marketer who plans to increase use of LinkedIn?
d. Why is a marketer who plans to increase use of LinkedIn and is a B2C marketer a joint event?

4.11 Referring to the contingency table in Problem 4.10, if a marketer is selected at random, what is the probability that
a. he or she plans to increase use of LinkedIn?
b. he or she is a B2C marketer?
c. he or she plans to increase use of LinkedIn *or* is a B2C marketer?
d. Explain the difference in the results in (b) and (c).

SELF Test **4.12** What business and technical skills are critical for today's business intelligence/analytics and information management professionals? As part of InformationWeek's 2013 U.S. IT Salary Survey, business intelligence/analytics and information management professionals, both staff and managers, were asked to indicate what business and technical skills are critical to their job. The list of business and technical skills included *Analyzing Data*. The following table summarizes the responses to this skill:

	PROFESSIONAL POSITION		
ANALYZING DATA	**Staff**	**Management**	**Total**
Critical	4,374	3,633	8,007
Not critical	3,436	2,631	6,067
Total	7,810	6,264	14,074

Source: Data extracted from "IT Salaries Show Slow Growth," *InformationWeek Reports*, April 2013, p. 40, **ubm.io/1ewjKT5**.

If a professional is selected at random, what is the probability that he or she

a. indicates analyzing data as critical to his or her job?
b. is a manager?
c. indicates analyzing data as critical to his or her job *or* is a manager?
d. Explain the difference in the results in (b) and (c).

4.13 Do Americans prefer Coke or Pepsi? A survey was conducted by Public Policy Polling (PPP) in 2013; the results were as follows:

	GENDER		
PREFERENCE	**Female**	**Male**	**Total**
Coke	120	95	215
Pepsi	95	80	175
Neither/Unsure	65	45	110
Total	280	220	500

Source: Data extracted from "Public Policy Polling" Report 2013, **bit.ly/YKXfzN**.

If an American is selected at random, what is the probability that he or she

a. prefers Pepsi?
b. is male *and* prefers Pepsi?
c. is male *or* prefers Pepsi?
d. Explain the difference in the results in (b) and (c).

4.14 A survey of 1,085 adults asked, "Do you enjoy shopping for clothing for yourself?" The results (data extracted from "Split Decision on Clothes Shopping," *USA Today*, January 28, 2011, p. 1B) indicated that 51% of the females enjoyed shopping for clothing for themselves as compared to 44% of the males. The sample sizes of males and females were not provided. Suppose that the results indicated that of 542 males, 238 answered yes. Of 543 females, 276 answered yes. Construct a contingency table to evaluate the probabilities. What is the probability that a respondent chosen at random

a. enjoys shopping for clothing for himself or herself?
b. is a female *and* enjoys shopping for clothing for herself?
c. is a female *or* is a person who enjoys shopping for clothing?
d. is a male *or* a female?

4.15 Each year, ratings are compiled concerning the performance of new cars during the first 90 days of use. Suppose that the cars have been categorized according to whether a car needs warranty-related repair (yes or no) and the country in which the company manufacturing a car is based (United States or not United States). Based on the data collected, the probability that the new car needs a warranty repair is 0.04, the probability that the car was manufactured by a U.S.-based company is 0.60, and the probability that the new car needs a warranty repair *and* was manufactured by a U.S.-based company is 0.025. Construct a contingency table to evaluate the probabilities of a warranty-related repair. What is the probability that a new car selected at random

a. needs a warranty repair?
b. needs a warranty repair *and* was manufactured by a U.S.-based company?
c. needs a warranty repair *or* was manufactured by a U.S.-based company?
d. needs a warranty repair *or* was not manufactured by a U.S.-based company?

4.2 Conditional Probability

Each example in Section 4.1 involves finding the probability of an event when sampling from the entire sample space. How do you determine the probability of an event if you know certain information about the events involved?

Computing Conditional Probabilities

Conditional probability refers to the probability of event *A*, given information about the occurrence of another event, *B*.

If this rule holds for two events, A and B, then A and B are independent. Therefore, there are two ways to determine independence:

1. Events A and B are independent if, and only if, $P(A|B) = P(A)$.
2. Events A and B are independent if, and only if, $P(A \text{ and } B) = P(A)P(B)$.

Marginal Probability Using the General Multiplication Rule

In Section 4.1, marginal probability was defined using Equation (4.2) on page 157. You can state the equation for marginal probability by using the general multiplication rule. If

$$P(A) = P(A \text{ and } B_1) + P(A \text{ and } B_2) + \cdots + P(A \text{ and } B_k)$$

then, using the general multiplication rule, Equation (4.8) defines the marginal probability.

MARGINAL PROBABILITY USING THE GENERAL MULTIPLICATION RULE

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_k)P(B_k) \quad \textbf{(4.8)}$$

where $B_1, B_2, \ldots, B_k$ are k mutually exclusive and collectively exhaustive events.

To illustrate Equation (4.8), refer to Table 4.1 on page 154. Let

$$\begin{aligned} P(A) &= \text{probability of planned to purchase} \\ P(B_1) &= \text{probability of actually purchased} \\ P(B_2) &= \text{probability of did not actually purchase} \end{aligned}$$

Then, using Equation (4.8), the probability of planned to purchase is

$$\begin{aligned} P(A) &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) \\ &= \left(\frac{200}{300}\right)\left(\frac{300}{1{,}000}\right) + \left(\frac{50}{700}\right)\left(\frac{700}{1{,}000}\right) \\ &= \frac{200}{1{,}000} + \frac{50}{1{,}000} = \frac{250}{1{,}000} = 0.25 \end{aligned}$$

Problems for Section 4.2

LEARNING THE BASICS

4.16 Consider the following contingency table:

	B	**B′**
A	10	20
A′	20	40

What is the probability of

a. $A|B$?
b. $A|B'$?
c. $A'|B'$?
d. Are events A and B independent?

4.17 Consider the following contingency table:

	B	**B′**
A	10	30
A′	25	35

What is the probability of

a. $A|B$?
b. $A'|B'$?
c. $A|B'$?
d. Are events A and B independent?

4.18 If $P(A \text{ and } B) = 0.4$ and $P(B) = 0.8$, find $P(A|B)$.

4.19 If $P(A) = 0.7$, $P(B) = 0.6$, and A and B are independent, find $P(A \text{ and } B)$.

4.20 If $P(A) = 0.3$, $P(B) = 0.4$, and $P(A \text{ and } B) = 0.2$, are A and B independent?

APPLYING THE CONCEPTS

4.21 Do males or females feel more tense or stressed out at work? A survey of employed adults conducted online by Harris Interactive on behalf of the American Psychological Association revealed the following:

	FELT TENSE OR STRESSED OUT AT WORK	
GENDER	Yes	No
Male	244	495
Female	282	480

Source: Data extracted from "The 2013 Work and Well-Being Survey," American Psychological Association and Harris Interactive, March 2013, p. 5, **bit.ly/11JGcPf**.

a. Given that the employed adult felt tense or stressed out at work, what is the probability that the employed adult was a male?
b. Given that the employed adult is male, what is the probability that he felt tense or stressed out at work?
c. Explain the difference in the results in (a) and (b).
d. Is feeling tense or stressed out at work and gender independent?

4.22 How will marketers change their social media use in the near future? A survey by Social Media Examiner reported that 76% of B2B marketers (marketers that focus primarily on attracting businesses) plan to increase their use of LinkedIn, as compared to 55% of B2C marketers (marketers that primarily target consumers). The survey was based on 1,945 B2B marketers and 1,868 B2C marketers. The following table summarizes the results:

	BUSINESS FOCUS		
INCREASE USE OF LINKEDIN?	B2B	B2C	Total
Yes	1,478	1,027	2,505
No	467	841	1,308
Total	1,945	1,868	3,813

Source: Data extracted from "2012 Social Media Marketing Industry Report," April 2012, p. 27, **bit.ly/HaWwDu**.

a. Suppose you know that the marketer is a B2B marketer. What is the probability that he or she plans to increase use of LinkedIn?
b. Suppose you know that the marketer is a B2C marketer. What is the probability that he or she plans to increase use of LinkedIn?
c. Are the two events, increase use of LinkedIn and business focus, independent? Explain.

4.23 Do Americans prefer Coke or Pepsi? A survey was conducted by Public Policy Polling (PPP) in 2013; the results were as follows:

	GENDER		
PREFERENCE	Female	Male	Total
Coke	120	95	215
Pepsi	95	80	175
Neither/Unsure	65	45	110
Total	280	220	500

Source: Data extracted from "Public Policy Polling" Report 2013, **bit.ly/YKXfzN**.

a. Given that an American is a male, what is the probability that he prefers Pepsi?
b. Given that an American is a female, what is the probability that she prefers Pepsi?
c. Is preference independent of gender? Explain.

SELF Test **4.24** What business and technical skills are critical for today's business intelligence/analytics and information management professionals? As part of InformationWeek's 2013 U.S. IT Salary Survey, business intelligence/analytics and information management professionals, both staff and managers, were asked to indicate what business and technical skills are critical to their job. The list of business and technical skills included *Analyzing Data*. The following table summarizes the responses to this skill:

	PROFESSIONAL POSITION		
ANALYZING DATA	Staff	Management	Total
Critical	4,374	3,633	8,007
Not critical	3,436	2,631	6,067
Total	7,810	6,264	14,074

Source: Data extracted from "IT Salaries Show Slow Growth," *InformationWeek Reports*, April 2013, p. 40, **ubm.io/1ewjKT5.**

a. Given that a professional is staff, what is the probability that the professional indicates analyzing data as critical to his or her job?
b. Given that a professional is staff, what is the probability that the professional does not indicate analyzing data as critical to his or her job?
c. Given that a professional is a manager, what is the probability that the professional indicates analyzing data as critical to his or her job?
d. Given that a professional is a manager, what is the probability that the professional does not indicate analyzing data as critical to his or her job?

4.25 A survey of 1,085 adults asked, "Do you enjoy shopping for clothing for yourself?" The results (data extracted from "Split Decision on Clothes Shopping," *USA Today*, January 28, 2011, p. 1B) indicated that 51% of the females enjoyed shopping for clothing for themselves as compared to 44% of the males. The sample sizes of males and females were not provided. Suppose that the results were as shown in the following table:

ENJOYS SHOPPING FOR CLOTHING	GENDER		
	Male	Female	Total
Yes	238	276	514
No	304	267	571
Total	542	543	1,085

a. Suppose that the respondent chosen is a female. What is the probability that she does not enjoy shopping for clothing?
b. Suppose that the respondent chosen enjoys shopping for clothing. What is the probability that the individual is a male?
c. Are enjoying shopping for clothing and the gender of the individual independent? Explain.

4.26 Each year, ratings are compiled concerning the performance of new cars during the first 90 days of use. Suppose that the cars have been categorized according to whether a car needs warranty-related repair (yes or no) and the country in which the company manufacturing a car is based (United States or not United States). Based on the data collected, the probability that the new car needs a warranty repair is 0.04, the probability that the car is manufactured by a U.S.-based company is 0.60, and the probability that the new car needs a warranty repair *and* was manufactured by a U.S.-based company is 0.025.
a. Suppose you know that a company based in the United States manufactured a particular car. What is the probability that the car needs a warranty repair?
b. Suppose you know that a company based in the United States did not manufacture a particular car. What is the probability that the car needs a warranty repair?
c. Are need for a warranty repair and location of the company manufacturing the car independent?

4.27 In 40 of the 62 years from 1950 through 2012 (in 2011 there was virtually no change), the S&P 500 finished higher after the first five days of trading. In 35 of those 40 years, the S&P 500 finished higher for the year. Is a good first week a good omen for the upcoming year? The following table gives the first-week and annual performance over this 62-year period:

FIRST WEEK	S&P 500'S ANNUAL PERFORMANCE	
	Higher	Lower
Higher	35	5
Lower	11	11

a. If a year is selected at random, what is the probability that the S&P 500 finished higher for the year?
b. Given that the S&P 500 finished higher after the first five days of trading, what is the probability that it finished higher for the year?
c. Are the two events "first-week performance" and "annual performance" independent? Explain.
d. Look up the performance after the first five days of 2013 and the 2013 annual performance of the S&P 500 at **finance.yahoo.com**. Comment on the results.

4.28 A standard deck of cards is being used to play a game. There are four suits (hearts, diamonds, clubs, and spades), each having 13 faces (ace, 2, 3, 4, 5, 6, 7, 8, 9, 10, jack, queen, and king), making a total of 52 cards. This complete deck is thoroughly mixed, and you will receive the first 2 cards from the deck, without replacement (the first card is not returned to the deck after it is selected).
a. What is the probability that both cards are queens?
b. What is the probability that the first card is a 10 and the second card is a 5 or 6?
c. If you were sampling with replacement (the first card is returned to the deck after it is selected), what would be the answer in (a)?
d. In the game of blackjack, the face cards (jack, queen, king) count as 10 points, and the ace counts as either 1 or 11 points. All other cards are counted at their face value. Blackjack is achieved if 2 cards total 21 points. What is the probability of getting blackjack in this problem?

4.29 A box of nine gloves contains two left-handed gloves and seven right-handed gloves.
a. If two gloves are randomly selected from the box, without replacement (the first glove is not returned to the box after it is selected), what is the probability that both gloves selected will be right-handed?
b. If two gloves are randomly selected from the box, without replacement (the first glove is not returned to the box after it is selected), what is the probability that there will be one right-handed glove and one left-handed glove selected?
c. If three gloves are selected, with replacement (the gloves are returned to the box after they are selected), what is the probability that all three will be left-handed?
d. If you were sampling with replacement (the first glove is returned to the box after it is selected), what would be the answers to (a) and (b)?

4.3 Bayes' Theorem

Bayes' theorem is used to revise previously calculated probabilities based on new information. Developed by Thomas Bayes in the eighteenth century (see references 1 and 6), Bayes' theorem is an extension of what you previously learned about conditional probability.

You can apply Bayes' theorem to the situation in which M&R Electronics World is considering marketing a new model of televisions. In the past, 40% of the new-model televisions have been successful, and 60% have been unsuccessful. Before introducing the new-model the word as Vi@gr@ or V1agra. What they overlooked was that the misspelled variants were even *more likely* to be found in a spam message than the original word. Thus, the misspelled variants made the job of spotting spam *easier* for the Bayesian filters.

Other spammers tried to fool the filters by adding "good" words, words that would have a low probability of being found in a spam message, or "rare" words, words not frequently encountered in any message. But these spammers overlooked the fact that the conditional probabilities are constantly updated and that words once considered "good" would be soon discarded from the good list by the filter as their $P(A|B)$, value increased. Likewise, as "rare" words grew more common in spam and yet stayed rare in ham, such words acted like the misspelled variants that others had tried earlier.

Even then, and perhaps after reading about Bayesian statistics, spammers thought that they could "break" Bayesian filters by inserting random words in their messages. Those random words would affect the filter by causing it to see many words whose $P(A|B)$, value would be low. The Bayesian filter would begin to label many spam messages as ham and end up being of no practical use. Spammers again overlooked that conditional probabilities are constantly updated.

Other spammers decided to eliminate all or most of the words in their messages and replace them with graphics so that Bayesian filters would have very few words with which to form conditional probabilities. But this approach failed, too, as Bayesian filters were rewritten to consider things other than words in a message. After all, Bayes' theorem concerns *events*, and "graphics present with no text" is as valid an event as "some word, *X*, present in a message." Other future tricks will ultimately fail for the same reason. (By the way, spam filters use non-Bayesian techniques as well, which makes spammers' lives even more difficult.)

Bayesian spam filters are an example of the unexpected way that applications of statistics can show up in your daily life. You will discover more examples as you read the rest of this book. By the way, the author of the two essays mentioned earlier was Thomas Bayes, who is a lot more famous for the second essay than the first essay, a failed attempt to use mathematics and logic to prove the existence of God.

Problems for Section 4.3

LEARNING THE BASICS

4.30 If $P(B) = 0.05$, $P(A|B) = 0.80$, $P(B') = 0.95$, and $P(A|B') = 0.40$, find $P(B|A)$.

4.31 If $P(B) = 0.30$, $P(A|B) = 0.60$, $P(B') = 0.70$, and and $P(A|B') = 0.50$, find $P(B|A)$.

APPLYING THE CONCEPTS

4.32 In Example 4.10 on page 171, suppose that the probability that a medical diagnostic test will give a positive result if the disease is not present is reduced from 0.02 to 0.01.

a. If the medical diagnostic test has given a positive result (indicating that the disease is present), what is the probability that the disease is actually present?

b. If the medical diagnostic test has given a negative result (indicating that the disease is not present), what is the probability that the disease is not present?

4.33 An advertising executive is studying television viewing habits of married men and women during prime-time hours. Based on past viewing records, the executive has determined that during prime time, husbands are watching television 60% of the time. When the husband is watching television, 40% of the time the wife is also watching. When the husband is not watching television, 30% of the time the wife is watching television.

a. Find the probability that if the wife is watching television, the husband is also watching television.

b. Find the probability that the wife is watching television during prime time.

SELF Test **4.34** Olive Construction Company is determining whether it should submit a bid for a new shopping center. In the past, Olive's main competitor, Base Construction Company, has submitted bids 70% of the time. If Base Construction Company does not bid on a job, the probability that Olive Construction Company will get the job is 0.50. If Base Construction Company bids on a job, the probability that Olive Construction Company will get the job is 0.25.

a. If Olive Construction Company gets the job, what is the probability that Base Construction Company did not bid?

b. What is the probability that Olive Construction Company will get the job?

4.35 Laid-off workers who become entrepreneurs because they cannot find meaningful employment with another company are known as *entrepreneurs by necessity*. *The Wall Street Journal* reported that these entrepreneurs by necessity are less likely to grow into large businesses than are *entrepreneurs by choice*. (Source: J. Bailey, "Desire—More Than Need—Builds a Business," *The Wall Street Journal*, May 21, 2001, p. B4.) This article states that 89% of the entrepreneurs in the United States are entrepreneurs by choice and 11% are entrepreneurs by necessity. Only 2% of entrepreneurs by necessity expect their new business to employ 20 or more people within five years, whereas 14% of entrepreneurs by choice expect to employ at least 20 people within five years.

a. If an entrepreneur is selected at random and that individual expects that his or her new business will employ 20 or more people within five years, what is the probability that this individual is an entrepreneur by choice?

b. Discuss several possible reasons why entrepreneurs by choice are more likely than entrepreneurs by necessity to believe that they will grow their businesses.

4.36 The editor of a textbook publishing company is trying to decide whether to publish a proposed business statistics textbook. Information on previous textbooks published indicates that 10% are huge successes, 20% are modest successes, 40% break even, and 30% are losers. However, before a publishing decision is made, the book will be reviewed. In the past, 99% of the huge successes received favorable reviews, 70% of the moderate successes received favorable reviews, 40% of the break-even books received favorable reviews, and 20% of the losers received favorable reviews.

a. If the proposed textbook receives a favorable review, how should the editor revise the probabilities of the various outcomes to take this information into account?

b. What proportion of textbooks receive favorable reviews?

4.37 A municipal bond service has three rating categories (A, B, and C). Suppose that in the past year, of the municipal bonds issued throughout the United States, 70% were rated A, 20% were rated B, and 10% were rated C. Of the municipal bonds rated A, 50% were issued by cities, 40% by suburbs, and 10% by rural areas. Of the municipal bonds rated B, 60% were issued by cities, 20% by suburbs, and 20% by rural areas. Of the municipal bonds rated C, 90% were issued by cities, 5% by suburbs, and 5% by rural areas.

a. If a new municipal bond is to be issued by a city, what is the probability that it will receive an A rating?

b. What proportion of municipal bonds are issued by cities?

c. What proportion of municipal bonds are issued by suburbs?

4.4 Counting Rules

In Equation (4.1) on page 152, the probability of occurrence of an outcome was defined as the number of ways the outcome occurs, divided by the total number of possible outcomes. Often, there are a large number of possible outcomes, and determining the exact number can be difficult. In such circumstances, rules have been developed for counting the number of possible outcomes. This section presents five different counting rules.

Counting Rule 1 Counting rule 1 determines the number of possible outcomes for a set of mutually exclusive and collectively exhaustive events.

COUNTING RULE 1

If any one of k different mutually exclusive and collectively exhaustive events can occur on each of n trials, the number of possible outcomes is equal to

$$k^n \tag{4.10}$$

For example, using Equation (4.10), the number of different possible outcomes from tossing a two-sided coin five times is $2^5 = 2 \times 2 \times 2 \times 2 \times 2 = 32$.

EXAMPLE 4.11

Rolling a Die Twice

Suppose you roll a die twice. How many different possible outcomes can occur?

SOLUTION If a six-sided die is rolled twice, using Equation (4.10), the number of different outcomes is $6^2 = 36$.

Counting Rule 2 The second counting rule is a more general version of the first counting rule and allows the number of possible events to differ from trial to trial.

COUNTING RULE 2

If there are k_1 events on the first trial, k_2 events on the second trial, . . . , and k_n events on the nth trial, then the number of possible outcomes is

$$(k_1)(k_2) \ldots (k_n) \tag{4.11}$$

For example, a state motor vehicle department would like to know how many license plate numbers are available if a license plate number consists of three letters followed by three numbers (0 through 9). Using Equation (4.11), if a license plate number consists of three letters followed by three numbers, the total number of possible outcomes is $(26)(26)(26)(10)(10)(10) = 17{,}576{,}000$.

EXAMPLE 4.14

Using Counting Rule 4

Modifying Example 4.13, if you have six books, but there is room for only four books on the shelf, in how many ways can you arrange these books on the shelf?

SOLUTION Using Equation (4.13), the number of ordered arrangements of four books selected from six books is equal to

$$_nP_x = \frac{n!}{(n-x)!} = \frac{6!}{(6-4)!} = \frac{(6)(5)(4)(3)(2)(1)}{(2)(1)} = 360$$

Counting Rule 5 In many situations, you are not interested in the *order* of the outcomes but only in the number of ways that x items can be selected from n items, *irrespective of order*. Each possible selection is called a **combination**.

COUNTING RULE 5: COMBINATIONS

The number of ways of selecting x objects from n objects, irrespective of order, is equal to

$$_nC_x = \frac{n!}{x!(n-x)!} \qquad \textbf{(4.14)}$$

where

n = total number of objects
x = number of objects to be arranged
$n!$ = n factorial = $n(n-1)\ldots(1)$
C = symbol for combinations[2]

[2]On many scientific calculators, there is a button labeled nCr that allows you to compute combinations. The symbol r is used instead of x.

If you compare this rule to counting rule 4, you see that it differs only in the inclusion of a term $x!$ in the denominator. When permutations were used, all of the arrangements of the x objects are distinguishable. With combinations, the $x!$ possible arrangements of objects are irrelevant.

EXAMPLE 4.15

Using Counting Rule 5

Modifying Example 4.14, if the order of the books on the shelf is irrelevant, in how many ways can you arrange these books on the shelf?

SOLUTION Using Equation (4.14), the number of combinations of four books selected from six books is equal to

$$_nC_x = \frac{n!}{x!(n-x)!} = \frac{6!}{4!(6-4)!} = \frac{(6)(5)(4)(3)(2)(1)}{(4)(3)(2)(1)(2)(1)} = 15$$

Problems for Section 4.4

APPLYING THE CONCEPTS

SELF Test **4.38** If there are 10 multiple-choice questions on an exam, each having three possible answers, how many different sequences of answers are there?

4.39 A lock on a bank vault consists of three dials, each with 30 positions. In order for the vault to open, each of the three dials must be in the correct position.

a. How many different possible dial combinations are there for this lock?

b. What is the probability that if you randomly select a position on each dial, you will be able to open the bank vault?

c. Explain why "dial combinations" are not mathematical combinations expressed by Equation (4.14).

4.40 a. If a coin is tossed seven times, how many different outcomes are possible?

b. If a die is tossed seven times, how many different outcomes are possible?

c. Discuss the differences in your answers to (a) and (b).

4.41 A particular brand of women's jeans is available in seven different sizes, three different colors, and three different styles. How many different women's jeans does the store manager need to order to have one pair of each type?

4.42 You would like to make a salad that consists of lettuce, tomato, cucumber, and peppers. You go to the supermarket, intending to purchase one variety of each of these ingredients. You discover that there are eight varieties of lettuce, four varieties of tomatoes, three varieties of cucumbers, and three varieties of peppers for sale at the supermarket. If you buy them all, how many different salads can you make?

4.43 A team is being formed that includes four different people. There are four different positions on the teams. How many different ways are there to assign the four people to the four positions?

4.44 In Major League Baseball, there are five teams in the Eastern Division of the National League: Atlanta, Florida, New York, Philadelphia, and Washington. How many different orders of finish are there for these five teams? (Assume that there are no ties in the standings.) Do you believe that all these orders are equally likely? Discuss.

4.45 Referring to Problem 4.44, how many different orders of finish are possible for the first four positions?

4.46 A gardener has six rows available in his vegetable garden to place tomatoes, eggplant, peppers, cucumbers, beans, and lettuce. Each vegetable will be allowed one and only one row. How many ways are there to position these vegetables in this garden?

4.47 There are eight members of a team. How many ways are there to select a team leader, assistant team leader, and team coordinator?

4.48 Four members of a group of 10 people are to be selected to a team. How many ways are there to select these four members?

4.49 A student has seven books that she would like to place in her backpack. However, there is room for only four books. Regardless of the arrangement, how many ways are there of placing four books into the backpack?

4.50 A daily lottery is conducted in which 2 winning numbers are selected out of 100 numbers. How many different combinations of winning numbers are possible?

4.51 A reading list for a course contains 20 articles. How many ways are there to choose 3 articles from this list?

4.5 Ethical Issues and Probability

Ethical issues can arise when any statements related to probability are presented to the public, particularly when these statements are part of an advertising campaign for a product or service. Unfortunately, many people are not comfortable with numerical concepts (see reference 5) and tend to misinterpret the meaning of the probability. In some instances, the misinterpretation is not intentional, but in other cases, advertisements may unethically try to mislead potential customers.

One example of a potentially unethical application of probability relates to advertisements for state lotteries. When purchasing a lottery ticket, the customer selects a set of numbers (such as 6) from a larger list of numbers (such as 54). Although virtually all participants know that they are unlikely to win the lottery, they also have very little idea of how unlikely it is for them to select all 6 winning numbers from the list of 54 numbers. They have even less of an idea of the probability of not selecting any winning numbers.

Given this background, you might consider a recent commercial for a state lottery that stated, "We won't stop until we have made everyone a millionaire" to be deceptive and possibly unethical. Do you think the state has any intention of ever stopping the lottery, given the fact that the state relies on it to bring millions of dollars into its treasury? Is it possible that the lottery can make everyone a millionaire? Is it ethical to suggest that the purpose of the lottery is to make everyone a millionaire?

Another example of a potentially unethical application of probability relates to an investment newsletter promising a 90% probability of a 20% annual return on investment. To make the claim in the newsletter an ethical one, the investment service needs to (a) explain the basis on which this probability estimate rests, (b) provide the probability statement in another format, such as 9 chances in 10, and (c) explain what happens to the investment in the 10% of the cases in which a 20% return is not achieved (e.g., is the entire investment lost?).

These are serious ethical issues. If you were going to write an advertisement for the state lottery that ethically describes the probability of winning a certain prize, what would you say? If you were going to write an advertisement for the investment newsletter that ethically states the probability of a 20% return on an investment, what would you say?

USING STATISTICS

Possibilities at M&R Electronics World, Revisited

Shock/Fotolia

As the marketing manager for M&R Electronics World, you analyzed the survey results of an intent-to-purchase study. This study asked the heads of 1,000 households about their intentions to purchase a large-screen HDTV sometime during the next 12 months, and as a follow-up, M&R surveyed the same people 12 months later to see whether such a television was purchased. In addition, for households purchasing large-screen HDTVs, the survey asked whether the television they purchased had a faster refresh rate, whether they also purchased a streaming media box in the past 12 months, and whether they were satisfied with their purchase of the large-screen HDTV.

By analyzing the results of these surveys, you were able to uncover many pieces of valuable information that will help you plan a marketing strategy to enhance sales and better target those households likely to purchase multiple or more expensive products. Whereas only 30% of the households actually purchased a large-screen HDTV, if a household indicated that it planned to purchase a large-screen HDTV in the next 12 months, there was an 80% chance that the household actually made the purchase. Thus the marketing strategy should target those households that have indicated an intention to purchase.

You determined that for households that purchased a television that had a faster refresh rate, there was a 47.5% chance that the household also purchased a streaming media box. You then compared this conditional probability to the marginal probability of purchasing a streaming media box, which was 36%. Thus, households that purchased televisions that had a faster refresh rate are more likely to purchase a streaming media box than are households that purchased large-screen HDTVs that have a standard refresh rate.

You were also able to apply Bayes' theorem to M&R Electronics World's market research reports. The reports investigate a potential new television model prior to its scheduled release. If a favorable report was received, then there was a 64% chance that the new television model would be successful. However, if an unfavorable report was received, there is only a 16% chance that the model would be successful. Therefore, the marketing strategy of M&R needs to pay close attention to whether a report's conclusion is favorable or unfavorable.

SUMMARY

This chapter began by developing the basic concepts of probability. You learned that probability is a numeric value from 0 to 1 that represents the chance, likelihood, or possibility that a particular event will occur. In addition to simple probability, you learned about conditional probabilities and independent events. Bayes' theorem was used to revise previously calculated probabilities based on new information. Throughout the chapter, contingency tables and decision trees were used to display information. You also learned about several counting rules. In the next chapter, important discrete probability distributions such as the binomial, Poisson, and hypergeometric distributions are developed.

REFERENCES

1. Bellhouse, D. R. "The Reverend Thomas Bayes, FRS: A Biography to Celebrate the Tercentenary of His Birth." *Statistical Science*, 19 (2004), 3–43.
2. Lowd, D., and C. Meek. "Good Word Attacks on Statistical Spam Filters." Presented at the Second Conference on Email and Anti-Spam, 2005.
3. *Microsoft Excel 2013*. Redmond, WA: Microsoft Corp., 2012.
4. *Minitab Release 16*. State College, PA: Minitab, Inc, 2010.
5. Paulos, J. A. *Innumeracy*. New York: Hill and Wang, 1988.
6. Silberman, S. "The Quest for Meaning," *Wired 8.02*, February 2000.
7. Zeller, T. "The Fight Against V1@gra (and Other Spam)." *The New York Times*, May 21, 2006, pp. B1, B6.

KEY EQUATIONS

Probability of Occurrence

$$\text{Probability of occurrence} = \frac{X}{T} \tag{4.1}$$

Marginal Probability

$$P(A) = P(A \text{ and } B_1) + P(A \text{ and } B_2) + \cdots + P(A \text{ and } B_k) \tag{4.2}$$

General Addition Rule

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) \tag{4.3}$$

Conditional Probability

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} \tag{4.4a}$$

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)} \tag{4.4b}$$

Independence

$$P(A|B) = P(A) \tag{4.5}$$

General Multiplication Rule

$$P(A \text{ and } B) = P(A|B)P(B) \tag{4.6}$$

Multiplication Rule for Independent Events

$$P(A \text{ and } B) = P(A)P(B) \tag{4.7}$$

Marginal Probability Using the General Multiplication Rule

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_k)P(B_k) \tag{4.8}$$

Bayes' Theorem

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_k)P(B_k)} \tag{4.9}$$

Counting Rule 1

$$k^n \tag{4.10}$$

Counting Rule 2

$$(k_1)(k_2) \ldots (k_n) \tag{4.11}$$

Counting Rule 3

$$n! = (n)(n-1) \ldots (1) \tag{4.12}$$

Counting Rule 4: Permutations

$${}_nP_x = \frac{n!}{(n-x)!} \tag{4.13}$$

Counting Rule 5: Combinations

$${}_nC_x = \frac{n!}{x!(n-x)!} \tag{4.14}$$

KEY TERMS

a priori probability 153
Bayes' theorem 169
certain event 152
collectively exhaustive 157
combination 176
complement 154
conditional probability 161
contingency table 155
decision tree 163
empirical probability 153
event 153
general addition rule 158
general multiplication rule 166
impossible event 152
independence 165
joint event 154
joint probability 156
marginal probability 157
multiplication rule for independent events 166
mutually exclusive 157
permutation 175
probability 152
sample space 154
simple event 153
simple probability 155
subjective probability 153
Venn diagram 155

CHECKING YOUR UNDERSTANDING

4.52 What are the differences between *a priori* probability, empirical probability, and subjective probability?

4.53 What is the difference between a simple event and a joint event?

4.54 How can you use the general addition rule to find the probability of occurrence of event *A* or *B*?

4.55 What is the difference between mutually exclusive events and collectively exhaustive events?

4.56 How does conditional probability relate to the concept of independence?

4.57 How does the multiplication rule differ for events that are and are not independent?

4.58 How can you use Bayes' theorem to revise probabilities in light of new information?

4.59 In Bayes' theorem, how does the prior probability differ from the revised probability?

CHAPTER REVIEW PROBLEMS

4.60 A survey by the Health Research Institute at PricewaterhouseCoopers LLP indicated that 80% of "young invincibles" (those aged 18 to 24) are likely to share health information through social media, as compared to 45% of "baby boomers" (those aged 45 to 64).

Source: Data extracted from "Social Media 'Likes' Healthcare: From Marketing to Social Business," Health Research Institute, April 2012, p. 8.

Suppose that the survey was based on 500 respondents from each of the two groups.

a. Construct a contingency table.
b. Give an example of a simple event and a joint event.
c. What is the probability that a randomly selected respondent is likely to share health information through social media?
d. What is the probability that a randomly selected respondent is likely to share health information through social media *and* is in the 45-to-64-year-old group?
e. Are the events "age group" and "likely to share health information through social media" independent? Explain.

4.61 SHL Americas provides a unique, global perspective of how talent is measured in its Global Assessment Trends Report. The report presents the results of an online survey conducted in late 2012 with HR professionals from companies headquartered throughout the world. The authors were interested in examining differences between respondents in *emerging economies* and those in *established economies* to provide relevant information for readers who may be creating assessment programs for organizations with global reach; one area of focus was on HR professionals' response to two statements: "My organization views HR as a strategic function" and "My organization uses talent information to make business decisions." The results are as follows:

	ORGANIZATION VIEWS HR AS A STRATEGIC FUNCTION		
ECONOMY	**Yes**	**No**	**Total**
Established	171	78	249
Emerging	222	121	343
Total	393	199	592

	ORGANIZATION USES INFORMATION TALENT TO MAKE BUSINESS DECISIONS		
ECONOMY	**Yes**	**No**	**Total**
Established	122	127	249
Emerging	130	213	343
Total	252	340	592

What is the probability that a randomly chosen HR professional
a. is from an established economy?
b. is from an established economy *or* agrees to the statement "My organization uses information talent to make business decisions?"
c. does not agree with the statement "My organization views HR as a strategic function" *and* is from an emerging economy?
d. does not agree with the statement "My organization views HR as a strategic function" *or* is from an emerging economy?
e. Suppose the randomly chosen HR professional does not agree with the statement "My organization views HR as a strategic function." What is the probability that the HR professional is from an emerging economy?
f. Are "My organization views HR as a strategic function" and the type of economy independent?
g. Is "My organization uses information talent to make business decisions" independent of the type of economy?

4.62 The 2012 Restaurant Industry Forecast takes a closer look at today's consumers. Based on a 2011 National Restaurant Association survey, consumers are divided into three segments (optimistic, cautious, and hunkered-down) based on their financial situation, current spending behavior, and economic outlook. Suppose the results, based on a sample of 100 males and 100 females, were as follows:

	GENDER		
CONSUMER SEGMENT	**Male**	**Female**	**Total**
Optimistic	26	16	42
Cautious	41	43	84
Hunkered-down	33	41	74
Total	100	100	200

Source: Data extracted from "The 2012 Restaurant Industry Forecast," National Restaurant Association, 2012, p. 12, **restaurant.org/research/forecast**.

If a consumer is selected at random, what is the probability that he or she

a. is classified as cautious?
b. is classified as optimistic or cautious?
c. is a male *or* is classified as hunkered-down?
d. is a male *and* is classified as hunkered-down?
e. Given that the consumer selected is a female, what is the probability that she is classified as optimistic?

4.63 A 2011 joint study by MIT Sloan Management Review and IBM Institute for Business Value reports a growing divide between those companies that are transforming themselves to take advantage of business analytics and those that have yet to embrace it. A survey of business executives, managers, and analysts from organizations around the world indicated that 31% of organizations are "aspirational users" (basic users of analytics), 45% are "experienced users" (moderate users of analytics), and 24% are "transformed users" (strong and sophisticated users of analytics). Furthermore, 62% of the transformed-user organizations indicated an intense level of focus on using analytics to better understand and connect with customers, as did 49% of the experienced-user organizations and 34% of the aspirational-user organizations. (Data extracted from "Analytics: The Widening Divide, How Companies Are Achieving Competitive Advantage Through Analytics," IBM Global Business Services, October, 2011.) If an organization is known to have an intense level of focus on using analytics to better understand and connect with customers, what is the probability that the organization is a transformed-user organization?

4.64 The CMO Council and SAS set out to better understand the key challenges, opportunities, and requirements that both chief marketing officers (CMOs) and chief information officers (CIOs) were facing in their journey to develop a more customer-centric enterprise. The following findings are from an online audit of 237 senior marketers and 210 senior IT executives. (Data extracted from "Big Data's Biggest Role: Aligning the CMO & CIO," March 2013, **bit.ly/11z7uKW**.)

	BIG DATA IS CRITICAL TO EXECUTING A CUSTOMER-CENTRIC PROGRAM		
EXECUTIVE GROUP	**Yes**	**No**	**Total**
Marketing	95	142	237
IT	107	103	210
Total	202	245	447

	FUNCTIONAL SILOS BLOCK AGGREGATION OF CUSTOMER DATA THROUGHOUT THE ORGANIZATION		
EXECUTIVE GROUP	**Yes**	**No**	**Total**
Marketing	122	115	237
IT	95	115	210
Total	217	230	447

a. What is the probability that a randomly selected executive identifies Big Data as critical to executing a customer-centric program?
b. Given that a randomly selected executive is a senior marketing executive, what is the probability that the executive identifies Big Data as critical to executing a customer-centric program?
c. Given that a randomly selected executive is a senior IT executive, what is the probability that the executive identifies Big Data as critical to executing a customer-centric program?
d. What is the probability that a randomly selected executive identifies that functional silos block aggregation of customer data throughout the organization?
e. Given that a randomly selected executive is a senior marketing executive, what is the probability that the executive identifies that functional silos block aggregation of customer data throughout the organization?
f. Given that a randomly selected executive is a senior IT executive, what is the probability that the executive identifies that functional silos block aggregation of customer data throughout the organization?
g. Comment on the results in (a) through (f).

4.65 A 2013 Sage North America survey examined the "financial literacy" of small business owners. The study found that 23% of small business owners indicated concern about income tax compliance for their business; 41% of small business owners use accounting software, given that the small business owner indicated concern about income tax compliance for his or her business. Given that a small business owner did not indicate concern about income tax compliance for his or her business, 58% of small business owners use accounting software. (Data extracted from "Sage Financial Capability Survey: What Small Business Owners Don't Understand Could Be Holding Them Back," April 17, 2013, **http://bit.ly/Z3FAqx**.)

a. Use Bayes' theorem to find the probability that a small business owner uses accounting software, given that the small business owner indicated concern about income tax compliance for his or her business.
b. Compare the result in (a) to the probability that a small business owner uses accounting software and comment on whether small business owners who are concerned about income tax compliance for their business are generally more likely to use accounting software than small business owners who are not concerned about income tax compliance for their business.