CHAPTER 2

PROBABILITY

2.1 INTRODUCTION

Historically, the oldest way of defining probabilities, the **classical probability concept**, applies when all possible outcomes are equally likely, as is presumably the case in most games of chance. We can then say that *if there are N equally likely possibilities, of which one must occur and n are regarded as favorable, or as a "success," then the probability of a "success" is given by the ratio* $\frac{n}{N}$.

EXAMPLE 2.1

What is the probability of drawing an ace from an ordinary deck of 52 playing cards?

Solution Since there are $n = 4$ aces among the $N = 52$ cards, the probability of drawing an ace is $\frac{4}{52} = \frac{1}{13}$. (It is assumed, of course, that each card has the same chance of being drawn.) ■

Although equally likely possibilities are found mostly in games of chance, the classical probability concept applies also in a great variety of situations where gambling devices are used to make random selections—when office space is assigned to teaching assistants by lot, when some of the families in a township are chosen in such a way that each one has the same chance of being included in a sample study, when machine parts are chosen for inspection so that each part produced has the same chance of being selected, and so forth.

A major shortcoming of the classical probability concept is its limited applicability, for there are many situations in which the possibilities that arise cannot all be regarded as equally likely. This would be the case, for instance, if we are concerned with the question whether it will rain on a given day, if we are concerned with the outcome of an election, or if we are concerned with a person's recovery from a disease.

Among the various probability concepts, most widely held is the **frequency interpretation**, according to which *the probability of an event* (*outcome or happening*) is the proportion of the time that events of the same kind will occur in the long run. If we say that the probability is 0.84 that a jet from Los Angeles to San Francisco will arrive on time, we mean (in accordance with the frequency interpretation) that such flights arrive on time 84 percent of the time. Similarly, if the weather bureau predicts that there is a 30 percent chance for rain (that is, a probability of 0.30), this means that under the same weather conditions it will rain 30 percent of the time. More generally, we say that an event has a probability of, say, 0.90, in the same sense in which we might say that our car will start in cold weather 90 percent of the time. We cannot guarantee what will happen on any particular occasion—the car may start and then it may not—but if we kept records over a long period of time, we should find that the proportion of "successes" is very close to 0.90.

The approach to probability that we shall use in this chapter is the **axiomatic approach**, in which probabilities are defined as "mathematical objects" that behave according to certain well-defined rules. Then, any one of the preceding probability concepts, or interpretations, can be used in applications as long as it is consistent with these rules.

2.2 SAMPLE SPACES

Since all probabilities pertain to the occurrence or nonoccurrence of events, let us explain first what we mean here by *event* and by the related terms *experiment*, *outcome*, and *sample space*.

It is customary in statistics to refer to any process of observation or measurement as an **experiment**. In this sense, an experiment may consist of the simple process of checking whether a switch is turned on or off; it may consist of counting the imperfections in a piece of cloth; or it may consist of the very complicated process of determining the mass of an electron. The results one obtains from an experiment, whether they are instrument readings, counts, "yes" or "no" answers, or values obtained through extensive calculations, are called the **outcomes** of the experiment.

The set of all possible outcomes of an experiment is called the **sample space**, and it is usually denoted by the letter S. Each outcome in a sample space is called an **element** of the sample space or simply a **sample point**. If a sample space has a finite number of elements, we may list the elements in the usual set notation; for instance, the sample space for the possible outcomes of one flip of a coin may be written

$$S = \{\mathrm{H}, \mathrm{T}\}$$

where H and T stand for head and tail. Sample spaces with a large or infinite number of elements are best described by a statement or rule; for example, if the possible outcomes of an experiment are the set of automobiles equipped with citizen band radios, the sample space may be written

$$S = \{x|x \text{ is an automobile with a CB radio}\}$$

This is read "S is the set of all x such that x is an automobile with a CB radio." Similarly, if S is the set of odd positive integers, we write

$$S = \{2k + 1 | k = 0, 1, 2, \ldots\}$$

How we formulate the sample space for a given situation will depend on the problem at hand. If an experiment consists of one roll of a die and we are interested in which face is turned up, we would use the sample space

$$S_1 = \{1, 2, 3, 4, 5, 6\}$$

However, if we are interested only in whether the face turned up is even or odd, we would use the sample space

$$S_2 = \{\text{even}, \text{odd}\}$$

This demonstrates that different sample spaces may well be used to describe an experiment. In general, *it is desirable to use sample spaces whose elements cannot be divided (partitioned or separated) into more primitive or more elementary kinds of outcomes.* In other words, *it is preferable that an element of a sample space not represent two or more outcomes that are distinguishable in some way.* Thus, in the preceding illustration S_1 would be preferable to S_2.

EXAMPLE 2.2

Describe a sample space that might be appropriate for an experiment in which we roll a pair of dice, one red and one green.

Solution The sample space that provides the most information consists of the 36 points given by

$$S_1 = \{(x, y) | x = 1, 2, \ldots, 6;\ y = 1, 2, \ldots, 6\}$$

where x represents the number turned up by the red die and y represents the number turned up by the green die. A second sample space, adequate for most purposes (though less desirable in general as it provides less information), is given by

$$S_2 = \{2, 3, 4, \ldots, 12\}$$

where the elements are the totals of the numbers turned up by the two dice. ■

Sample spaces are usually classified according to the number of elements that they contain. In the preceding example the sample spaces S_1 and S_2 contained a **finite** number of elements; but if a coin is flipped until a head appears for the first time, this could happen on the first flip, the second flip, the third flip, the fourth flip, ..., and there are infinitely many possibilities. For this experiment we obtain the sample space

$$S = \{\text{H}, \text{TH}, \text{TTH}, \text{TTTH}, \text{TTTTH}, \ldots\}$$

with an unending sequence of elements. But even here the number of elements can be matched one-to-one with the whole numbers, and in this sense the sample space is said to be **countable**. If a sample space contains a finite number of elements or an infinite though countable number of elements, it is said to be **discrete**.

The outcomes of some experiments are neither finite nor countably infinite. Such is the case, for example, when one conducts an investigation to determine the distance that a certain make of car will travel over a prescribed test course on 5 liters of gasoline. If we assume that distance is a variable that can be measured to any desired degree of accuracy, there is an infinity of possibilities (distances) that cannot be matched one-to-one with the whole numbers. Also, if we want to measure the amount of time it takes for two chemicals to react, the amounts making up the sample space are infinite in number and not countable. Thus, sample spaces need not be discrete. If a sample space consists of a continuum, such as all the points of a line segment or all the points in a plane, it is said to be **continuous**. Continuous sample spaces arise in practice whenever the outcomes of experiments are measurements of physical properties, such as temperature, speed, pressure, length, ..., that are measured on continuous scales.

2.3 EVENTS

In many problems we are interested in results that are not given directly by a specific element of a sample space.

EXAMPLE 2.3

With reference to the first sample space S_1 on page 25, describe the event A that the number of points rolled with the die is divisible by 3.

Solution Among 1, 2, 3, 4, 5, and 6, only 3 and 6 are divisible by 3. Therefore, A is represented by the subset $\{3, 6\}$ of the sample space S_1. ■

EXAMPLE 2.4

With reference to the sample space S_1 of Example 2.2, describe the event B that the total number of points rolled with the pair of dice is 7.

Solution Among the 36 possibilities, only (1, 6), (2, 5), (3, 4), (4, 3), (5, 2), and (6, 1) yield a total of 7. So, we write

$$B = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$$

Note that in Figure 2.1 the event of rolling a total of 7 with the two dice is represented by the set of points inside the region bounded by the dotted line. ■

In the same way, any event (outcome or result) can be identified with a collection of points, which constitute a subset of an appropriate sample space. Such a subset consists of all the elements of the sample space for which the event occurs, and in probability and statistics we identify the subset with the event. Thus, by definition, an **event** is a subset of a sample space.

EXAMPLE 2.5

If someone takes three shots at a target and we care only whether each shot is a hit or a miss, describe a suitable sample space, the elements of the sample space that

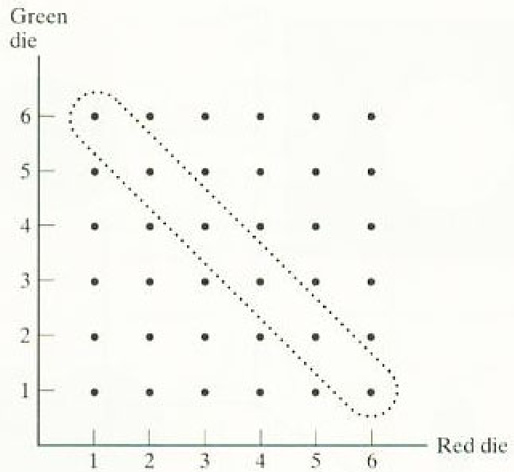

FIGURE 2.1: Rolling a total of 7 with a pair of dice.

constitute event M that the person will miss the target three times in a row, and the elements of event N that the person will hit the target once and miss it twice.

Solution If we let 0 and 1 represent a miss and a hit, respectively, the eight possibilities (0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 0), (1, 0, 1), (0, 1, 1), and (1, 1, 1) may be displayed as in Figure 2.2. Thus, it can be seen that

$$M = \{(0, 0, 0)\}$$

and

$$N = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$ ■

EXAMPLE 2.6

Construct a sample space for the length of the useful life of a certain electronic component and indicate the subset that represents the event F that the component fails before the end of the sixth year.

Solution If t is the length of the component's useful life in years, the sample space may be written $S = \{t | t \geqq 0\}$, and the subset $F = \{t | 0 \leqq t < 6\}$ is the event that the component fails before the end of the sixth year. ■

According to our definition, any event is a subset of an appropriate sample space, but it should be observed that the converse is not necessarily true. For discrete sample spaces, all subsets are events, but in the continuous case some rather abstruse point sets must be excluded for mathematical reasons. This is discussed further in some of the more advanced texts listed among the references at the end of this chapter, but it is of no consequence as far as the work of this book is concerned.

In many problems of probability we are interested in events that are actually combinations of two or more events, formed by taking **unions**, **intersections**, and

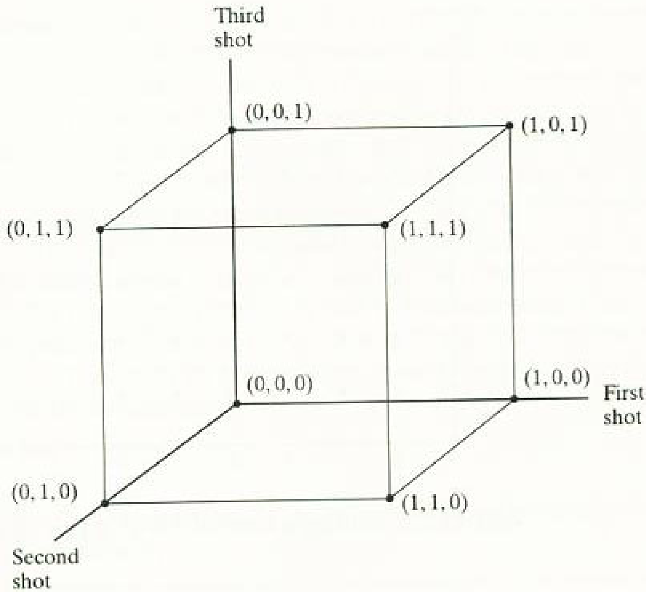

FIGURE 2.2: Sample space for Example 2.5.

complements. Although the reader must surely be familiar with these terms, let us review briefly that, if A and B are any two subsets of a sample space S, their union $A \cup B$ is the subset of S that contains all the elements that are either in A, in B, or in both; their intersection $A \cap B$ is the subset of S that contains all the elements that are in both A and B; and the complement A' of A is the subset of S that contains all the elements of S that are not in A. Some of the rules that control the formation of unions, intersections, and complements may be found in Exercises 2.1 through 2.4.

Sample spaces and events, particularly relationships among events, are often depicted by means of **Venn diagrams**, in which the sample space is represented by a rectangle, while events are represented by regions within the rectangle, usually by circles or parts of circles. For instance, the shaded regions of the four Venn diagrams of Figure 2.3 represent, respectively, event A, the complement of event A, the union of events A and B, and the intersection of events A and B. When we are dealing with three events, we usually draw the circles as in Figure 2.4. Here, the regions are numbered 1 through 8 for easy reference.

To indicate special relationships among events, we sometimes draw diagrams like those of Figure 2.5. Here, the one on the left serves to indicate that events A and B are **mutually exclusive**; that is, the two sets have no elements in common (or the two events cannot both occur). When A and B are mutually exclusive, we write $A \cap B = \emptyset$, where $\emptyset$ denotes the **empty set**, which has no elements at all. The diagram on the right serves to indicate that A is contained in B, and symbolically we express this by writing $A \subset B$.

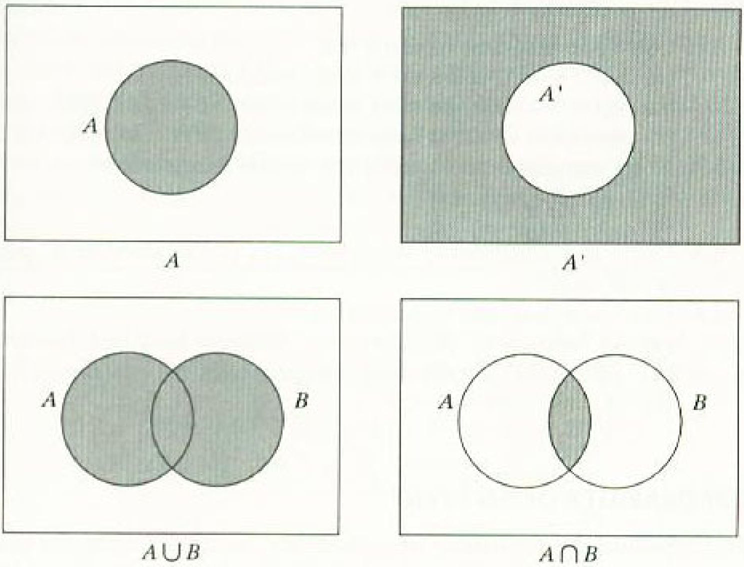

FIGURE 2.3: Venn diagrams.

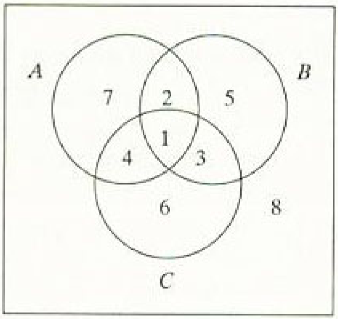

FIGURE 2.4: Venn diagram.

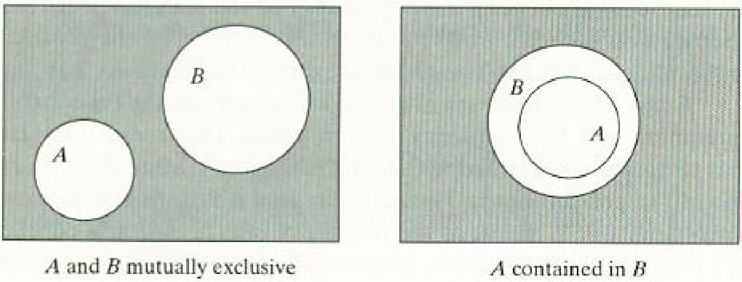

FIGURE 2.5: Diagrams showing special relationships among events.

EXERCISES

2.1. Use Venn diagrams to verify that

(a) $(A \cup B) \cup C$ is the same event as $A \cup (B \cup C)$;

(b) $A \cap (B \cup C)$ is the same event as $(A \cap B) \cup (A \cap C)$;

(c) $A \cup (B \cap C)$ is the same event as $(A \cup B) \cap (A \cup C)$.

2.2. Use Venn diagrams to verify the two **De Morgan laws**:

(a) $(A \cap B)' = A' \cup B'$;

(b) $(A \cup B)' = A' \cap B'$.

2.3. Use Venn diagrams to verify that if A is contained in B, then $A \cap B = A$ and $A \cap B' = \emptyset$.

2.4. Use Venn diagrams to verify that

(a) $(A \cap B) \cup (A \cap B') = A$;

(b) $(A \cap B) \cup (A \cap B') \cup (A' \cap B) = A \cup B$;

(c) $A \cup (A' \cap B) = A \cup B$.

2.4 THE PROBABILITY OF AN EVENT

To formulate the postulates of probability, we shall follow the practice of denoting events by means of capital letters, and we shall write the probability of event A as $P(A)$, the probability of event B as $P(B)$, and so forth. The following postulates of probability apply only to discrete sample spaces, S.

POSTULATE 1 The probability of an event is a nonnegative real number; that is, $P(A) \geqq 0$ for any subset A of S.

POSTULATE 2 $P(S) = 1$.

POSTULATE 3 If $A_1, A_2, A_3, \ldots$, is a finite or infinite sequence of mutually exclusive events of S, then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$$

Postulates per se require no proof, but if the resulting theory is to be applied, we must show that the postulates are satisfied when we give probabilities a "real" meaning. Let us illustrate this here in connection with the frequency interpretation; the relationship between the postulates and the classical probability concept will be discussed on page 33, while the relationship between the postulates and subjective probabilities is left for the reader to examine in Exercises 2.16 and 2.82.

Since proportions are always positive or zero, the first postulate is in complete agreement with the frequency interpretation. The second postulate states indirectly that certainty is identified with a probability of 1; after all, it is always assumed that one of the possibilities in S must occur, and it is to this certain event that we assign a probability of 1. As far as the frequency interpretation is concerned, a probability of 1 implies that the event in question will occur 100 percent of the time or, in other words, that it is certain to occur.

Taking the third postulate in the simplest case, that is, for two mutually exclusive events A_1 and A_2, it can easily be seen that it is satisfied by the frequency interpretation. If one event occurs, say, 28 percent of the time, another event occurs 39 percent of the time, and the two events cannot both occur at the same time (that is, they are mutually exclusive), then one or the other will occur $28+39 = 67$ percent of the time. Thus, the third postulate is satisfied, and the same kind of argument applies when there are more than two mutually exclusive events.

Before we study some of the immediate consequences of the postulates of probability, let us emphasize the point that the three postulates do not tell us how to assign probabilities to events; they merely restrict the ways in which it can be done.

EXAMPLE 2.7

An experiment has four possible outcomes, A, B, C, and D, that are mutually exclusive. Explain why the following assignments of probabilities are not permissible:

(a) $P(A) = 0.12$, $P(B) = 0.63$, $P(C) = 0.45$, $P(D) = -0.20$;
(b) $P(A) = \frac{9}{120}$, $P(B) = \frac{45}{120}$, $P(C) = \frac{27}{120}$, $P(D) = \frac{46}{120}$.

Solution

(a) $P(D) = -0.20$ violates Postulate 1;
(b) $P(S) = P(A \cup B \cup C \cup D) = \frac{9}{120} + \frac{45}{120} + \frac{27}{120} + \frac{46}{120} = \frac{127}{120} \neq 1$, and this violates Postulate 2. ■

Of course, in actual practice probabilities are assigned on the basis of past experience, on the basis of a careful analysis of all underlying conditions, on the basis of subjective judgments, or on the basis of assumptions—sometimes the assumption that all possible outcomes are equiprobable.

To assign a probability measure to a sample space, it is not necessary to specify the probability of each possible subset. This is fortunate, for a sample space with as few as 20 possible outcomes has already $2^{20} = 1{,}048{,}576$ subsets [the general formula follows directly from part (a) of Exercise 1.14], and the number of subsets grows very rapidly when there are 50 possible outcomes, 100 possible outcomes, or more. Instead of listing the probabilities of all possible subsets, we often list the probabilities of the individual outcomes, or sample points of S, and then make use of the following theorem.

THEOREM 2.1. If A is an event in a discrete sample space S, then $P(A)$ equals the sum of the probabilities of the individual outcomes comprising A.

Proof. Let $O_1, O_2, O_3, \ldots$, be the finite or infinite sequence of outcomes that comprise the event A. Thus,

$$A = O_1 \cup O_2 \cup O_3 \cdots$$

and since the individual outcomes, the O's, are mutually exclusive, the third postulate of probability yields

$$P(A) = P(O_1) + P(O_2) + P(O_3) + \cdots$$

This completes the proof. □

To use this theorem, we must be able to assign probabilities to the individual outcomes of experiments. How this is done in some special situations is illustrated by the following examples.

EXAMPLE 2.8

If we twice flip a balanced coin, what is the probability of getting at least one head?

Solution The sample space is $S = \{HH, HT, TH, TT\}$, where H and T denote head and tail. Since we assume that the coin is balanced, these outcomes are equally likely and we assign to each sample point the probability $\frac{1}{4}$. Letting A denote the event that we will get at least one head, we get $A = \{HH, HT, TH\}$ and

$$\begin{aligned} P(A) &= P(\text{HH}) + P(\text{HT}) + P(\text{TH}) \\ &= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} \\ &= \frac{3}{4} \end{aligned}$$

■

EXAMPLE 2.9

A die is loaded in such a way that each odd number is twice as likely to occur as each even number. Find $P(G)$, where G is the event that a number greater than 3 occurs on a single roll of the die.

Solution The sample space is $S = \{1, 2, 3, 4, 5, 6\}$. Hence, if we assign probability w to each even number and probability $2w$ to each odd number, we find that $2w + w + 2w + w + 2w + w = 9w = 1$ in accordance with Postulate 2. It follows that $w = \frac{1}{9}$ and

$$P(G) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}$$

■

If a sample space is countably infinite, probabilities will have to be assigned to the individual outcomes by means of a mathematical rule, preferably by means of a formula or equation.

EXAMPLE 2.10

If, for a given experiment, $O_1, O_2, O_3, \ldots$, is an infinite sequence of outcomes, verify that

$$P(O_i) = \left(\frac{1}{2}\right)^i \qquad \text{for } i = 1, 2, 3, \ldots$$

is, indeed, a probability measure.

Solution Since the probabilities are all positive, it remains to be shown that $P(S) = 1$. Getting

$$P(S) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots$$

and making use of the formula for the sum of the terms of an infinite geometric progression, we find that

$$P(S) = \frac{\frac{1}{2}}{1 - \frac{1}{2}} = 1 \qquad \blacksquare$$

In connection with the preceding example, the word "sum" in Theorem 2.1 will have to be interpreted so that it includes the value of an infinite series.

As we shall see in Chapter 5, the probability measure of Example 2.10 would be appropriate, for example, if O_i is the event that a person flipping a balanced coin will get a tail for the first time on the ith flip of the coin. Thus, the probability that the first tail will come on the third, fourth, or fifth flip of the coin is

$$\left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^5 = \frac{7}{32}$$

and the probability that the first tail will come on an odd-numbered flip of the coin is

$$\left(\frac{1}{2}\right)^1 + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^5 + \cdots = \frac{\frac{1}{2}}{1 - \frac{1}{4}} = \frac{2}{3}$$

Here again we made use of the formula for the sum of the terms of an infinite geometric progression.

If an experiment is such that we can assume equal probabilities for all the sample points, as was the case in Example 2.8, we can take advantage of the following special case of Theorem 2.1.

Theorem 2.2. If an experiment can result in any one of N different equally likely outcomes, and if n of these outcomes together constitute event A, then the probability of event A is

$$P(A) = \frac{n}{N}$$

Proof. Let $O_1, O_2, \ldots, O_N$ represent the individual outcomes in S, each with probability $\frac{1}{N}$. If A is the union of n of these mutually exclusive outcomes, and it does not matter which ones, then

$$\begin{aligned} P(A) &= P(O_1 \cup O_2 \cup \cdots \cup O_n) \\ &= P(O_1) + P(O_2) + \cdots + P(O_n) \\ &= \underbrace{\frac{1}{N} + \frac{1}{N} + \cdots + \frac{1}{N}}_{n \text{ terms}} \\ &= \frac{n}{N} \qquad \square \end{aligned}$$

Observe that the formula $P(A) = \frac{n}{N}$ of Theorem 2.2 is identical with the one for the classical probability concept (see page 23). Indeed, what we have shown here is that the classical probability concept is consistent with the postulates of probability—it follows from the postulates in the special case where the individual outcomes are all equiprobable.

EXAMPLE 2.11

A five-card poker hand dealt from a deck of 52 playing cards is said to be a full house if it consists of three of a kind and a pair. If all the five-card hands are equally likely, what is the probability of being dealt a full house?

Solution The number of ways in which we can be dealt a particular full house, say three kings and two aces, is $\binom{4}{3}\binom{4}{2}$. Since there are 13 ways of selecting the face value for the three of a kind and for each of these there are 12 ways of selecting the face value for the pair, there are altogether

$$n = 13 \cdot 12 \cdot \binom{4}{3}\binom{4}{2}$$

different full houses. Also, the total number of equally likely five-card poker hands is

$$N = \binom{52}{5}$$

and it follows by Theorem 2.2 that the probability of getting a full house is

$$P(A) = \frac{n}{N} = \frac{13 \cdot 12 \binom{4}{3}\binom{4}{2}}{\binom{52}{5}} = 0.0014$$

■

2.5 SOME RULES OF PROBABILITY

Based on the three postulates of probability, we can derive many other rules that have important applications. Among them, the next four theorems are immediate consequences of the postulates.

Theorem 2.3. If A and A' are complementary events in a sample space S, then

$$P(A') = 1 - P(A)$$

Proof. In the second and third steps of the proof that follows, we make use of the definition of a complement, according to which A and A' are mutually exclusive and $A \cup A' = S$. Thus, we write

$$\begin{aligned} 1 &= P(S) && \text{(by Postulate 2)} \\ &= P(A \cup A') \\ &= P(A) + P(A') && \text{(by Postulate 3)} \end{aligned}$$

and it follows that $P(A') = 1 - P(A)$. □

In connection with the frequency interpretation, this result implies that if an event occurs, say, 37 percent of the time, then it does not occur 63 percent of the time.

THEOREM 2.4. $P(\emptyset) = 0$ for any sample space S.

Proof. Since S and $\emptyset$ are mutually exclusive and $S \cup \emptyset = S$ in accordance with the definition of the empty set $\emptyset$, it follows that

$$\begin{aligned} P(S) &= P(S \cup \emptyset) \\ &= P(S) + P(\emptyset) \qquad \text{(by Postulate 3)} \end{aligned}$$

and, hence, that $P(\emptyset) = 0$. □

It is important to note that it does not necessarily follow from $P(A) = 0$ that $A = \emptyset$. In practice, we often assign 0 probability to events that, in colloquial terms, would not happen in a million years. For instance, there is the classical example that we assign a probability of 0 to the event that a monkey set loose on a typewriter will type Plato's *Republic* word for word without a mistake. As we shall see in Chapters 3 and 6, the fact that $P(A) = 0$ does not imply that $A = \emptyset$ is of relevance, especially, in the continuous case.

THEOREM 2.5. If A and B are events in a sample space S and $A \subset B$, then $P(A) \leqq P(B)$.

Proof. Since $A \subset B$, we can write

$$B = A \cup (A' \cap B)$$

as can easily be verified by means of a Venn diagram. Then, since A and $A' \cap B$ are mutually exclusive, we get

$$\begin{aligned} P(B) &= P(A) + P(A' \cap B) \qquad \text{(by Postulate 3)} \\ &\geqq P(A) \qquad \text{(by Postulate 1)} \end{aligned}$$

□

In words, this theorem states that if A is a subset of B, then $P(A)$ cannot be greater than $P(B)$. For instance, the probability of drawing a heart from an ordinary deck of 52 playing cards cannot be greater than the probability of drawing a red card. Indeed, the probability is $\frac{1}{4}$, compared with $\frac{1}{2}$.

THEOREM 2.6. $0 \leqq P(A) \leqq 1$ for any event A.

Proof. Using Theorem 2.5 and the fact that $\emptyset \subset A \subset S$ for any event A in S, we have

$$P(\emptyset) \leqq P(A) \leqq P(S)$$

Then, $P(\emptyset) = 0$ and $P(S) = 1$ leads to the result that

$$0 \leqq P(A) \leqq 1$$

□

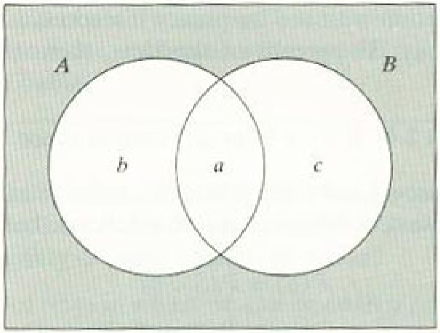

FIGURE 2.6: Venn diagram for proof of Theorem 2.7.

The third postulate of probability is sometimes referred to as the **special addition rule**; it is special in the sense that events $A_1, A_2, A_3, \ldots$, must all be mutually exclusive. For any two events A and B, there exists the **general addition rule:**

THEOREM 2.7. If A and B are any two events in a sample space S, then

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

Proof. Assigning the probabilities a, b, and c to the mutually exclusive events $A \cap B$, $A \cap B'$, and $A' \cap B$ as in the Venn diagram of Figure 2.6, we find that

$$\begin{aligned} P(A \cup B) &= a + b + c \\ &= (a + b) + (c + a) - a \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$ □

EXAMPLE 2.12

In a large metropolitan area, the probabilities are 0.86, 0.35, and 0.29 that a family (randomly chosen for a sample survey) owns a color television set, a HDTV set, or both kinds of sets. What is the probability that a family owns either or both kinds of sets?

Solution If A is the event that a family in this metropolitan area owns a color television set and B is the event that it owns a HDTV set, we have $P(A) = 0.86$, $P(B) = 0.35$, and $P(A \cap B) = 0.29$; substitution into the formula of Theorem 2.7 yields

$$\begin{aligned} P(A \cup B) &= 0.86 + 0.35 - 0.29 \\ &= 0.92 \end{aligned}$$ ■

EXAMPLE 2.13

Near a certain exit of I-17, the probabilities are 0.23 and 0.24 that a truck stopped at a roadblock will have faulty brakes or badly worn tires. Also, the probability is

0.38 that a truck stopped at the roadblock will have faulty brakes and/or badly worn tires. What is the probability that a truck stopped at this roadblock will have faulty brakes as well as badly worn tires?

Solution If B is the event that a truck stopped at the roadblock will have faulty brakes and T is the event that it will have badly worn tires, we have $P(B) = 0.23$, $P(T) = 0.24$, and $P(B \cup T) = 0.38$; substitution into the formula of Theorem 2.7 yields

$$0.38 = 0.23 + 0.24 - P(B \cap T)$$

Solving for $P(B \cap T)$, we thus get

$$P(B \cap T) = 0.23 + 0.24 - 0.38 = 0.09$$ ■

Repeatedly using the formula of Theorem 2.7, we can generalize this addition rule so that it will apply to any number of events. For instance, for three events we get

Theorem 2.8. If A, B, and C are any three events in a sample space S, then

$$\begin{aligned} P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) \\ - P(B \cap C) + P(A \cap B \cap C) \end{aligned}$$

Proof. Writing $A \cup B \cup C$ as $A \cup (B \cup C)$ and using the formula of Theorem 2.7 twice, once for $P[A \cup (B \cup C)]$ and once for $P(B \cup C)$, we get

$$\begin{aligned} P(A \cup B \cup C) &= P[A \cup (B \cup C)] \\ &= P(A) + P(B \cup C) - P[A \cap (B \cup C)] \\ &= P(A) + P(B) + P(C) - P(B \cap C) \\ &\quad - P[A \cap (B \cup C)] \end{aligned}$$

Then, using the distributive law that the reader was asked to verify in part (b) of Exercise 2.1, we find that

$$\begin{aligned} P[A \cap (B \cup C)] &= P[(A \cap B) \cup (A \cap C)] \\ &= P(A \cap B) + P(A \cap C) - P[(A \cap B) \cap (A \cap C)] \\ &= P(A \cap B) + P(A \cap C) - P(A \cap B \cap C) \end{aligned}$$

and hence that

$$\begin{aligned} P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) \\ - P(B \cap C) + P(A \cap B \cap C) \end{aligned}$$ □

(In Exercise 2.12 the reader will be asked to give an alternative proof of this theorem, based on the method used in the text to prove Theorem 2.7.)

EXAMPLE 2.14

If a person visits his dentist, suppose that the probability that he will have his teeth cleaned is 0.44, the probability that he will have a cavity filled is 0.24, the probability that he will have a tooth extracted is 0.21, the probability that he will have his teeth cleaned and a cavity filled is 0.08, the probability that he will have his teeth cleaned and a tooth extracted is 0.11, the probability that he will have a cavity filled and a tooth extracted is 0.07, and the probability that he will have his teeth cleaned, a cavity filled, and a tooth extracted is 0.03. What is the probability that a person visiting his dentist will have at least one of these things done to him?

Solution If C is the event that the person will have his teeth cleaned, F is the event that he will have a cavity filled, and E is the event that he will have a tooth extracted, we are given $P(C) = 0.44$, $P(F) = 0.24$, $P(E) = 0.21$, $P(C \cap F) = 0.08$, $P(C \cap E) = 0.11$, $P(F \cap E) = 0.07$, and $P(C \cap F \cap E) = 0.03$, and substitution into the formula of Theorem 2.8 yields

$$\begin{aligned} P(C \cup F \cup E) &= 0.44 + 0.24 + 0.21 - 0.08 - 0.11 - 0.07 + 0.03 \\ &= 0.66 \end{aligned}$$

■

EXERCISES

2.5. Use parts (a) and (b) of Exercise 2.4 to show that
(a) $P(A) \geqq P(A \cap B)$;
(b) $P(A) \leqq P(A \cup B)$.

2.6. Referring to Figure 2.6, verify that

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

2.7. Referring to Figure 2.6 and letting $P(A' \cap B') = d$, verify that

$$P(A' \cap B') = 1 - P(A) - P(B) + P(A \cap B)$$

2.8. The event that "A or B but not both" will occur can be written as

$$(A \cap B') \cup (A' \cap B)$$

Express the probability of this event in terms of $P(A)$, $P(B)$, and $P(A \cap B)$.

2.9. Use the formula of Theorem 2.7 to show that
(a) $P(A \cap B) \leqq P(A) + P(B)$;
(b) $P(A \cap B) \geqq P(A) + P(B) - 1$.

2.10. Use the Venn diagram of Figure 2.7 with the probabilities a, b, c, d, e, f, and g assigned to $A \cap B \cap C$, $A \cap B \cap C'$, ..., and $A \cap B' \cap C'$ to show that if $P(A) = P(B) = P(C) = 1$, then $P(A \cap B \cap C) = 1$. [*Hint*: Start with the argument that since $P(A) = 1$, it follows that $e = c = f = 0$.]

2.11. Give an alternative proof of Theorem 2.7 by making use of the relationships $A \cup B = A \cup (A' \cap B)$ and $B = (A \cap B) \cup (A' \cap B)$.

2.12. Use the Venn diagram of Figure 2.7 and the method by which we proved Theorem 2.7 to prove Theorem 2.8.

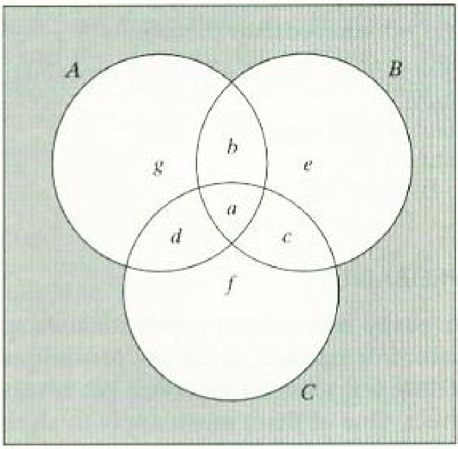

FIGURE 2.7: Venn diagram for Exercises 2.10, 2.12, and 2.13.

2.13. Duplicate the method of proof used in Exercise 2.12 to show that

$$\begin{aligned} P(A \cup B \cup C \cup D) = {} & P(A) + P(B) + P(C) + P(D) - P(A \cap B) \\ & - P(A \cap C) - P(A \cap D) - P(B \cap C) - P(B \cap D) \\ & - P(C \cap D) + P(A \cap B \cap C) + P(A \cap B \cap D) \\ & + P(A \cap C \cap D) + P(B \cap C \cap D) \\ & - P(A \cap B \cap C \cap D) \end{aligned}$$

(*Hint:* With reference to the Venn diagram of Figure 2.7, divide each of the eight regions into two parts, designating one to be inside D and the other outside D and letting $a, b, c, d, e, f, g, h, i, j, k, l, m, n, o$, and p be the probabilities associated with the resulting 16 regions.)

2.14. Prove by induction that

$$P(E_1 \cup E_2 \cup \cdots \cup E_n) \leqq \sum_{i=1}^{n} P(E_i)$$

for any finite sequence of events $E_1, E_2, \ldots$, and E_n.

2.15. The **odds** that an event will occur are given by the ratio of the probability that the event will occur to the probability that it will not occur, provided neither probability is zero. Odds are usually quoted in terms of positive integers having no common factor. Show that if the odds are A to B that an event will occur, its probability is

$$p = \frac{a}{a+b}$$

2.16. Subjective probabilities may be determined by exposing persons to risk-taking situations and finding the odds at which they would consider it fair to bet on the outcome. The odds are then converted into probabilities by means of the formula of Exercise 2.15. For instance, if a person feels that 3 to 2 are fair odds that a business venture will succeed (or that it would be fair to bet \$30 against

\$20 that it will succeed), the probability is $\frac{3}{3+2} = 0.6$ that the business venture will succeed. Show that if subjective probabilities are determined in this way, they satisfy

(a) Postulate 1 on page 30;

(b) Postulate 2.

See also Exercise 2.82.

2.6 CONDITIONAL PROBABILITY

Difficulties can easily arise when probabilities are quoted without specification of the sample space. For instance, if we ask for the probability that a lawyer makes more than \$50,000 per year, we may well get several different answers, and they may all be correct. One of them might apply to all those who are actively engaged in the practice of law, and so forth. Since the choice of the sample space (that is, the set of all possibilities under consideration) is by no means always self-evident, it often helps to use the symbol $P(A|S)$ to denote the **conditional probability** of event A relative to the sample space S or, as we also call it, "the probability of A given S." The symbol $P(A|S)$ makes it explicit that we are referring to a particular sample space S, and it is preferable to the abbreviated notation $P(A)$ unless the tacit choice of S is clearly understood. It is also preferable when we want to refer to several sample spaces in the same example. If A is the event that a person makes more than \$50,000 per year, G is the event that a person is a law school graduate, L is the event that a person is licensed to practice law, and E is the event that a person is actively engaged in the practice of law, then $P(A|G)$ is the probability that a law school graduate makes more than \$50,000 per year, $P(A|L)$ is the probability that a person licensed to practice law makes more than \$50,000 per year, and $P(A|E)$ is the probability that a person actively engaged in the practice of law makes more than \$50,000 per year.

Some ideas connected with conditional probabilities are illustrated in the following example.

EXAMPLE 2.15

A consumer research organization has studied the services under warranty provided by the 50 new-car dealers in a certain city, and its findings are summarized in the following table.

	Good service under warranty	*Poor service under warranty*
In business 10 years or more	16	4
In business less than 10 years	10	20

If a person randomly selects one of these new-car dealers, what is the probability that he gets one who provides good service under warranty? Also, if a person randomly selects one of the dealers who has been in business for 10 years or more, what is the probability that he gets one who provides good service under warranty?

Solution By "randomly" we mean that, in each case, all possible selections are equally likely, and we can therefore use the formula of Theorem 2.2. If we let G denote the selection of a dealer who provides good service under warranty, and if we let $n(G)$ denote the number of elements in G and $n(S)$ the number of elements in the whole sample space, we get

$$P(G) = \frac{n(G)}{n(S)} = \frac{16+10}{50} = 0.52$$

This answers the first question.

For the second question, we limit ourselves to the reduced sample space, which consists of the first line of the table, that is, the $16 + 4 = 20$ dealers who have been in business 10 years or more. Of these, 16 provide good service under warranty, and we get

$$P(G|T) = \frac{16}{20} = 0.80$$

where T denotes the selection of a dealer who has been in business 10 years or more. This answers the second question and, as should have been expected, $P(G|T)$ is considerably higher than $P(G)$. ■

Since the numerator of $P(G|T)$ is $n(T \cap G) = 16$ in the preceding example, the number of dealers who have been in business for 10 years or more and provide good service under warranty, and the denominator is $n(T)$, the number of dealers who have been in business 10 years or more, we can write symbolically

$$P(G|T) = \frac{n(T \cap G)}{n(T)}$$

Then, if we divide the numerator and the denominator by $n(S)$, the total number of new-car dealers in the given city, we get

$$P(G|T) = \frac{\frac{n(T \cap G)}{n(S)}}{\frac{n(T)}{n(S)}} = \frac{P(T \cap G)}{P(T)}$$

and we have, thus, expressed the conditional probability $P(G|T)$ in terms of two probabilities defined for the whole sample space S.

Generalizing from the preceding, let us now make the following definition of conditional probability.

Definition 2.1. If A and B are any two events in a sample space S and $P(A) \neq 0$, the **conditional probability** of B given A is

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

EXAMPLE 2.16

With reference to Example 2.15, what is the probability that one of the dealers who has been in business less than 10 years will provide good service under warranty?

Solution Since $P(T' \cap G) = \frac{10}{50} = 0.20$ and $P(T') = \frac{10+20}{50} = 0.60$, substitution into the formula yields

$$P(G|T') = \frac{P(T' \cap G)}{P(T')} = \frac{0.20}{0.60} = \frac{1}{3}$$

■

Although we introduced the formula for $P(B|A)$ by means of an example in which the possibilities were all equally likely, this is not a requirement for its use.

EXAMPLE 2.17

With reference to the loaded die of Example 2.9, what is the probability that the number of points rolled is a perfect square? Also, what is the probability that it is a perfect square given that it is greater than 3?

Solution If A is the event that the number of points rolled is greater than 3 and B is the event that it is a perfect square, we have $A = \{4, 5, 6\}$, $B = \{1, 4\}$, and $A \cap B = \{4\}$. Since the probabilities of rolling a 1, 2, 3, 4, 5, or 6 with the die are $\frac{2}{9}, \frac{1}{9}, \frac{2}{9}, \frac{1}{9}, \frac{2}{9}$, and $\frac{1}{9}$ (see page 32), we find that the answer to the first question is

$$P(B) = \frac{2}{9} + \frac{1}{9} = \frac{1}{3}$$

To determine $P(B|A)$, we first calculate

$$P(A \cap B) = \frac{1}{9} \quad \text{and} \quad P(A) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}$$

Then, substituting into the formula of Definition 2.1, we get

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{9}}{\frac{4}{9}} = \frac{1}{4}$$

■

To verify that the formula of Definition 2.1 has yielded the "right" answer in the preceding example, we have only to assign probability v to the two even numbers in the reduced sample space A and probability $2v$ to the odd number, such that the sum of the three probabilities is equal to 1. We thus have $v + 2v + v = 1$, $v = \frac{1}{4}$, and, hence, $P(B|A) = \frac{1}{4}$ as before.

EXAMPLE 2.18

A manufacturer of airplane parts knows from past experience that the probability is 0.80 that an order will be ready for shipment on time, and it is 0.72 that an order will be ready for shipment on time and will also be delivered on time. What is the probability that such an order will be delivered on time given that it was ready for shipment on time?

Solution If we let R stand for the event that an order is ready for shipment on time and D be the event that it is delivered on time, we have $P(R) = 0.80$ and $P(R \cap D) = 0.72$, and it follows that

$$P(D|R) = \frac{P(R \cap D)}{P(R)} = \frac{0.72}{0.80} = 0.90$$

Thus, 90 percent of the shipments will be delivered on time provided they are shipped on time. Note that $P(R|D)$, the probability that a shipment that is delivered on time was also ready for shipment on time, cannot be determined without further information; for this purpose we would also have to know $P(D)$. ■

If we multiply the expressions on both sides of the formula of Definition 2.1 by $P(A)$, we obtain the following **multiplication rule**.

THEOREM 2.9. If A and B are any two events in a sample space S and $P(A) \neq 0$, then

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

In words, the probability that A and B will both occur is the product of the probability of A and the conditional probability of B given A. Alternatively, if $P(B) \neq 0$, the probability that A and B will both occur is the product of the probability of B and the conditional probability of A given B; symbolically,

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

To derive this alternative multiplication rule, we interchange A and B in the formula of Theorem 2.9 and make use of the fact that $A \cap B = B \cap A$.

EXAMPLE 2.19

If we randomly pick two television tubes in succession from a shipment of 240 television tubes of which 15 are defective, what is the probability that they will both be defective?

Solution If we assume equal probabilities for each selection (which its what we mean by "randomly" picking the tubes), the probability that the first tube will be defective is $\frac{15}{240}$, and the probability that the second tube will be defective given that the first tube is defective is $\frac{14}{239}$. Thus, the probability that both tubes will be defective is $\frac{15}{240} \cdot \frac{14}{239} = \frac{7}{1{,}912}$. This assumes that we are **sampling without replacement**; that is, the first tube is not replaced before the second tube is selected. ■

EXAMPLE 2.20

Find the probabilities of randomly drawing two aces in succession from an ordinary deck of 52 playing cards if we sample

(a) without replacement;

(b) with replacement.

Solution

(a) If the first card is not replaced before the second card is drawn, the probability of getting two aces in succession is

$$\frac{4}{52}\cdot\frac{3}{51}=\frac{1}{221}$$

(b) If the first card is replaced before the second card is drawn, the corresponding probability is

$$\frac{4}{52}\cdot\frac{4}{52}=\frac{1}{169}$$ ■

In the situations described in the two preceding examples there is a definite temporal order between the two events A and B. In general, this need not be the case when we write $P(A|B)$ or $P(B|A)$. For instance, we could ask for the probability that the first card drawn was an ace given that the second card drawn (without replacement) is an ace—the answer would also be $\frac{3}{51}$.

Theorem 2.9 can easily be generalized so that it applies to more than two events; for instance, for three events we have

THEOREM 2.10. If A, B, and C are any three events in a sample space S such that $P(A\cap B)\neq 0$, then

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

Proof. Writing $A\cap B\cap C$ as $(A\cap B)\cap C$ and using the formula of Theorem 2.9 twice, we get

$$\begin{aligned}P(A\cap B\cap C)&=P[(A\cap B)\cap C]\\&=P(A\cap B)\cdot P(C|A\cap B)\\&=P(A)\cdot P(B|A)\cdot P(C|A\cap B)\end{aligned}$$ □

EXAMPLE 2.21

A box of fuses contains 20 fuses, of which five are defective. If three of the fuses are selected at random and removed from the box in succession without replacement, what is the probability that all three fuses are defective?

Solution If A is the event that the first fuse is defective, B is the event that the second fuse is defective, and C is the event that the third fuse is defective, then $P(A)=\frac{5}{20}$, $P(B|A)=\frac{4}{19}$, $P(C|A\cap B)=\frac{3}{18}$, and substitution into the formula yields

$$\begin{aligned}P(A\cap B\cap C)&=\frac{5}{20}\cdot\frac{4}{19}\cdot\frac{3}{18}\\&=\frac{1}{114}\end{aligned}$$ ■

Further generalization of Theorems 2.9 and 2.10 to k events is straightforward, and the resulting formula can be proved by mathematical induction.

2.7 INDEPENDENT EVENTS

Informally speaking, two events A and B are **independent** if the occurrence or nonoccurrence of either one does not affect the probability of the occurrence of the other. For instance, in the preceding example the selections would all have been independent had each fuse been replaced before the next one was selected; the probability of getting a defective fuse would have remained $\frac{5}{20}$.

Symbolically, two events A and B are independent if $P(B|A) = P(B)$ and $P(A|B) = P(A)$, and it can be shown that either of these equalities implies the other when both of the conditional probabilities exist, that is, when neither $P(A)$ nor $P(B)$ equals zero (see Exercise 2.21).

Now, if we substitute $P(B)$ for $P(B|A)$ into the formula of Theorem 2.9, we get

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

and we shall use this as our formal definition of independence.

Definition 2.2. Two events A and B are **independent** if and only if

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

Reversing the steps, we can also show that Definition 2.2 implies the definition of independence that we gave earlier.

If two events are not independent, they are said to be **dependent**. In the derivation of the formula of Definition 2.2, we assume that $P(B|A)$ exists and, hence, that $P(A) \neq 0$. For mathematical convenience, we shall let the definition apply also when $P(A) = 0$ and/or $P(B) = 0$.

EXAMPLE 2.22

A coin is tossed three times and the eight possible outcomes, HHH, HHT, HTH, THH, HTT, THT, TTH, and TTT, are assumed to be equally likely. If A is the event that a head occurs on each of the first two tosses, B is the event that a tail occurs on the third toss, and C is the event that exactly two tails occur in the three tosses, show that

(a) events A and B are independent;

(b) events B and C are dependent.

Solution Since

$$\begin{aligned} A &= \{\text{HHH, HHT}\} \\ B &= \{\text{HHT, HTT, THT, TTT}\} \\ C &= \{\text{HTT, THT, TTH}\} \\ A \cap B &= \{\text{HHT}\} \\ B \cap C &= \{\text{HTT, THT}\} \end{aligned}$$

the assumption that the eight possible outcomes are all equiprobable yields $P(A) = \frac{1}{4}$, $P(B) = \frac{1}{2}$, $P(C) = \frac{3}{8}$, $P(A \cap B) = \frac{1}{8}$, and $P(B \cap C) = \frac{1}{4}$.

(a) Since $P(A) \cdot P(B) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8} = P(A \cap B)$, events A and B are independent.

(b) Since $P(B) \cdot P(C) = \frac{1}{2} \cdot \frac{3}{8} = \frac{3}{16} \neq P(B \cap C)$, events B and C are not independent. ■

In connection with Definition 2.2, it can be shown that if A and B are independent, then so are A and B', A' and B, and A' and B'. For instance,

Theorem 2.11. If A and B are independent, then A and B' are also independent.

Proof. Since $A = (A \cap B) \cup (A \cap B')$, as the reader was asked to show in part (a) of Exercise 2.4, $A \cap B$ and $A \cap B'$ are mutually exclusive, and A and B are independent by assumption, we have

$$\begin{aligned} P(A) &= P[(A \cap B) \cup (A \cap B')] \\ &= P(A \cap B) + P(A \cap B') \\ &= P(A) \cdot P(B) + P(A \cap B') \end{aligned}$$

It follows that

$$\begin{aligned} P(A \cap B') &= P(A) - P(A) \cdot P(B) \\ &= P(A) \cdot [1 - P(B)] \\ &= P(A) \cdot P(B') \end{aligned}$$

and hence that A and B' are independent. □

In Exercises 2.22 and 2.23 the reader will be asked to show that if A and B are independent, then A' and B are independent and so are A' and B', and if A and B are dependent, then A and B' are dependent.

To extend the concept of independence to more than two events, let us make the following definition.

Definition 2.3. Events $A_1, A_2, \ldots,$ and A_k are **independent** if and only if the probability of the intersection of any $2, 3, \ldots,$ or k of these events equals the product of their respective probabilities.

For three events A, B, and C, for example, independence requires that

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

and

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

It is of interest to note that three or more events can be **pairwise independent** without being independent.

EXAMPLE 2.23

Figure 2.8 shows a Venn diagram with probabilities assigned to its various regions. Verify that A and B are independent, A and C are independent, and B and C are independent, but A, B, and C are not independent.

Solution As can be seen from the diagram, $P(A) = P(B) = P(C) = \frac{1}{2}$, $P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4}$, and $P(A \cap B \cap C) = \frac{1}{4}$. Thus,

$$P(A) \cdot P(B) = \frac{1}{4} = P(A \cap B)$$

$$P(A) \cdot P(C) = \frac{1}{4} = P(A \cap C)$$

$$P(B) \cdot P(C) = \frac{1}{4} = P(B \cap C)$$

but

$$P(A) \cdot P(B) \cdot P(C) = \frac{1}{8} \neq P(A \cap B \cap C)$$ ■

Incidentally, the preceding example can be given a "real" interpretation by considering a large room that has three separate switches controlling the ceiling lights. These lights will be on when all three switches are "up" and hence also when one of the switches is "up" and the other two are "down." If A is the event that the first switch is "up," B is the event that the second switch is "up," and C is the event that the third switch is "up," the Venn diagram of Figure 2.8 shows a possible set of probabilities associated with the switches being "up" or "down" when the ceiling lights are on.

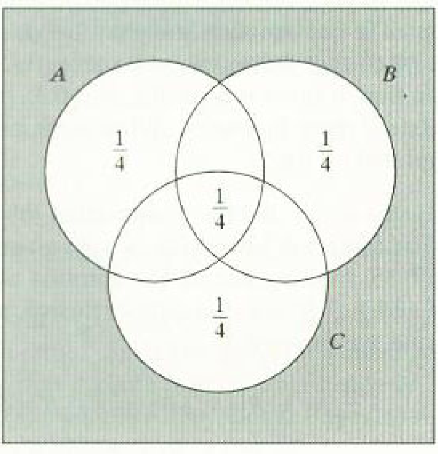

FIGURE 2.8: Venn diagram for Example 2.23.

It can also happen that $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$ without A, B, and C being pairwise independent—this the reader will be asked to verify in Exercise 2.24.

Of course, if we are given that certain events are independent, the probability that they will all occur is simply the product of their respective probabilities.

EXAMPLE 2.24

Find the probabilities of getting

(a) three heads in three random tosses of a balanced coin;
(b) four sixes and then another number in five random rolls of a balanced die.

Solution

(a) Multiplying the respective probabilities, we get

$$\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

(b) Multiplying the respective probabilities, we get

$$\frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{5}{7,776}$$

■

2.8 BAYES' THEOREM

In many situations the outcome of an experiment depends on what happens in various intermediate stages. The following is a simple example in which there is one intermediate stage consisting of two alternatives:

EXAMPLE 2.25

The completion of a construction job may be delayed because of a strike. The probabilities are 0.60 that there will be a strike, 0.85 that the construction job will be completed on time if there is no strike, and 0.35 that the construction job will be completed on time if there is a strike. What is the probability that the construction job will be completed on time?

Solution If A is the event that the construction job will be completed on time and B is the event that there will be a strike, we are given $P(B) = 0.60$, $P(A|B') = 0.85$, and $P(A|B) = 0.35$. Making use of the formula of part (a) of Exercise 2.4, the fact that $A \cap B$ and $A \cap B'$ are mutually exclusive, and the alternative form of the multiplication rule, we can write

$$\begin{aligned} P(A) &= P[(A \cap B) \cup (A \cap B')] \\ &= P(A \cap B) + P(A \cap B') \\ &= P(B) \cdot P(A|B) + P(B') \cdot P(A|B') \end{aligned}$$

Then, substituting the given numerical values, we get

$$P(A) = (0.60)(0.35) + (1 - 0.60)(0.85)$$
$$= 0.55$$

■

An immediate generalization of this kind of situation is the case where the intermediate stage permits k different alternatives (whose occurrence is denoted by $B_1, B_2, \ldots, B_k$). It requires the following theorem, sometimes called the **rule of total probability** or the **rule of elimination**.

THEOREM 2.12. If the events $B_1, B_2, \ldots,$ and B_k constitute a partition of the sample space S and $P(B_i) \neq 0$ for $i = 1, 2, \ldots, k$, then for any event A in S

$$P(A) = \sum_{i=1}^{k} P(B_i) \cdot P(A|B_i)$$

As was defined in the footnote on page 9, the B's constitute a partition of the sample space if they are pairwise mutually exclusive and if their union equals S. A formal proof of Theorem 2.12 consists, essentially, of the same steps we used in Example 2.25, and it is left to the reader in Exercise 2.32.

EXAMPLE 2.26

The members of a consulting firm rent cars from three rental agencies: 60 percent from agency 1, 30 percent from agency 2, and 10 percent from agency 3. If 9 percent of the cars from agency 1 need a tune-up, 20 percent of the cars from agency 2 need a tune-up, and 6 percent of the cars from agency 3 need a tune-up, what is the probability that a rental car delivered to the firm will need a tune-up?

Solution If A is the event that the car needs a tune-up, and B_1, B_2, and B_3 are the events that the car comes from rental agencies 1, 2, or 3, we have $P(B_1) = 0.60$, $P(B_2) = 0.30$, $P(B_3) = 0.10$, $P(A|B_1) = 0.09$, $P(A|B_2) = 0.20$, and $P(A|B_3) = 0.06$. Substituting these values into the formula of Theorem 2.12, we get

$$P(A) = (0.60)(0.09) + (0.30)(0.20) + (0.10)(0.06)$$
$$= 0.12$$

Thus, 12 percent of all the rental cars delivered to this firm will need a tune-up. ■

With reference to the preceding example, suppose that we are interested in the following question: If a rental car delivered to the consulting firm needs a tune-up, what is the probability that it came from rental agency 2? To answer questions of this kind, we need the following theorem, called **Bayes' theorem**:

THEOREM 2.13. If $B_1, B_2, \ldots, B_k$ constitute a partition of the sample space S and $P(B_i) \neq 0$ for $i = 1, 2, \ldots, k$, then for any event A in S such that

$P(A) \neq 0$

$$P(B_r|A) = \frac{P(B_r) \cdot P(A|B_r)}{\sum_{i=1}^{k} P(B_i) \cdot P(A|B_i)}$$

for $r = 1, 2, \ldots, k$.

In words, the probability that event A was reached via the rth branch of the tree diagram of Figure 2.9, given that it was reached via one of its k branches, is the *ratio* of the probability associated with the rth branch to the sum of the probabilities associated with all k branches of the tree.

Proof. Writing $P(B_r|A) = \frac{P(A \cap B_r)}{P(A)}$ in accordance with the definition of conditional probability, we have only to substitute $P(B_r) \cdot P(A|B_r)$ for $P(A \cap B_r)$ and the formula of Theorem 2.12 for $P(A)$. □

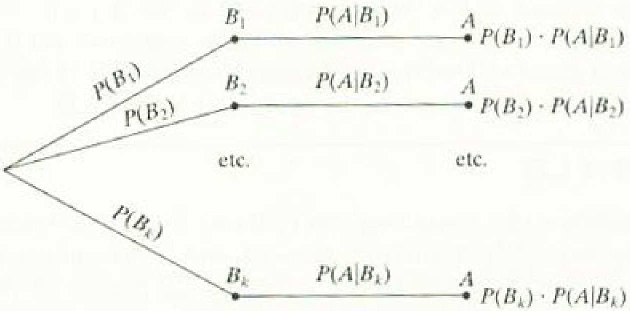

FIGURE 2.9: Tree diagram for Bayes' theorem.

EXAMPLE 2.27

With reference to Example 2.26, if a rental car delivered to the consulting firm needs a tune-up, what is the probability that it came from rental agency 2?

Solution Substituting the probabilities on page 49 into the formula of Theorem 2.13, we get

$$\begin{aligned} P(B_2|A) &= \frac{(0.30)(0.20)}{(0.60)(0.09) + (0.30)(0.20) + (0.10)(0.06)} \\ &= \frac{0.060}{0.120} \\ &= 0.5 \end{aligned}$$

Observe that although only 30 percent of the cars delivered to the firm come from agency 2, 50 percent of those requiring a tune-up come from that agency. ■

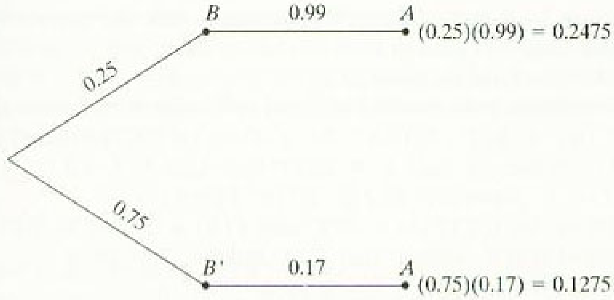

FIGURE 2.10: Tree diagram for Example 2.28.

EXAMPLE 2.28

In a certain state, 25 percent of all cars emit excessive amounts of pollutants. If the probability is 0.99 that a car emitting excessive amounts of pollutants will fail the state's vehicular emission test, and the probability is 0.17 that a car not emitting excessive amounts of pollutants will nevertheless fail the test, what is the probability that a car that fails the test actually emits excessive amounts of pollutants?

Solution Picturing this situation as in Figure 2.10, we find that the probabilities associated with the two branches of the tree diagram are $(0.25)(0.99) = 0.2475$ and $(1 - 0.25)(0.17) = 0.1275$. Thus, the probability that a car that fails the test actually emits excessive amounts of pollutants is

$$\frac{0.2475}{0.2475 + 0.1275} = 0.66$$

Of course, this result could also have been obtained without the diagram by substituting directly into the formula of Bayes' theorem. ■

Although Bayes' theorem follows from the postulates of probability and the definition of conditional probability, it has been the subject of extensive controversy. There can be no question about the validity of Bayes' theorem, but considerable arguments have been raised about the assignment of the **prior probabilities** $P(B_i)$. Also, a good deal of mysticism surrounds Bayes' theorem because it entails a "backward," or "inverse," sort of reasoning, that is, reasoning "from effect to cause." For instance, in Example 2.28, failing the test is the effect and emitting excessive amounts of pollutants is a possible cause.

EXERCISES

2.17. Show that the postulates of probability are satisfied by conditional probabilities. In other words, show that if $P(B) \neq 0$, then

(a) $P(A|B) \geqq 0$;

(b) $P(B|B) = 1$;

(c) $P(A_1 \cup A_2 \cup \ldots |B) = P(A_1|B) + P(A_2|B) + \cdots$ for any sequence of mutually exclusive events $A_1, A_2, \ldots$.

2.18. Show by means of numerical examples that $P(B|A) + P(B|A')$

(a) may be equal to 1;

(b) need not be equal to 1.

2.19. Duplicating the method of proof of Theorem 2.10, show that $P(A \cap B \cap C \cap D) = P(A) \cdot P(B|A) \cdot P(C|A \cap B) \cdot P(D|A \cap B \cap C)$ provided that $P(A \cap B \cap C) \neq 0$.

2.20. Given three events A, B, and C such that $P(A \cap B \cap C) \neq 0$ and $P(C|A \cap B) = P(C|B)$, show that $P(A|B \cap C) = P(A|B)$.

2.21. Show that if $P(B|A) = P(B)$ and $P(B) \neq 0$, then $P(A|B) = P(A)$.

2.22. Show that if events A and B are independent, then

(a) events A' and B are independent;

(b) events A' and B' are independent.

2.23. Show that if events A and B are dependent, then events A and B' are dependent.

2.24. Refer to Figure 2.11 to show that $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$ does not necessarily imply that A, B, and C are all pairwise independent.

2.25. Refer to Figure 2.11 to show that if A is independent of B and A is independent of C, then B is not necessarily independent of C.

2.26. Refer to Figure 2.11 to show that if A is independent of B and A is independent of C, then A is not necessarily independent of $B \cup C$.

2.27. If events A, B, and C are independent, show that

(a) A and $B \cap C$ are independent.

(b) A and $B \cup C$ are independent.

2.28. If $P(A|B) < P(A)$, prove that $P(B|A) < P(B)$.

2.29. If $A_1, A_2, \ldots, A_n$ are independent events, prove that

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = 1 - \{1 - P|A_1\} \cdot \{1 - P|A_2\} \ldots \{1 - P|A_n\}$$

2.30. Show that $2^k - k - 1$ conditions must be satisfied for k events to be independent.

2.31. For any event A, show that A and $\emptyset$ are independent.

2.32. Prove Theorem 2.12 by making use of the following generalization of the distributive law given in part (b) of Exercise 2.1:

$$A \cap (B_1 \cup B_2 \cup \cdots \cup B_k) = (A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_k)$$

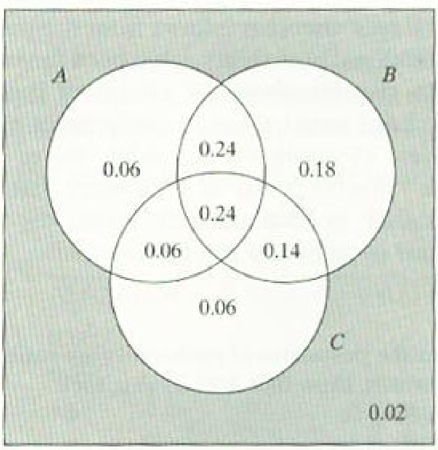

FIGURE 2.11: Diagram for Exercises 2.24, 2.25, and 2.26.

2.33. Suppose that a die has n sides numbered $i = 1, 2, \ldots n$. Assume that the probability of it coming up on the side numbered i is the same for each value of i. The die is rolled n times (assume independence) and a "match" is defined to be the occurrence of side i on the i^{th} roll. Prove that the probability of at least one match is given by

$$1 - \left(\frac{n-1}{n}\right)^n = 1 - \left(1 - \frac{1}{n}\right)^n$$

2.34. Show that $P(A \cup B) \geq 1 - P(A') - P(B')$ for any two events A and B defined in the sample space S. (*Hint*: Use Venn diagrams.)

2.9 THE THEORY IN PRACTICE

The word "probability" is a part of everyday language, but it is difficult to defined this word without using the word "probable" or its synonym "likely" in the definition. To illustrate, Webster's *Third New International Dictionary* defines "probability" as "the quality or state of being probable." If the concept of probability is to be used in mathematics and scientific applications, we require a more exact, less circular, definition.

The postulates of probability given in Section 2.4 satisfy this criterion. Together with the rules given in Section 2.5, this definition lends itself to calculations of probabilities that "make sense" and that can be verified experimentally. The entire theory of statistics is based on the notion of probability. It seems remarkable that the entire structure of probability, and therefore of statistics, can be built on the relatively straightforward foundation given in this chapter.

Probabilities were first considered in games of chance, or gambling. Players of various games of chance observed that there seemed to be "rules" that governed the roll of dice or the results of spinning a roulette wheel. Some of them went as far as to postulate some of these rules entirely on the basis of experience. But differences arose among gamblers about probabilities, and they brought their questions to the noted mathematicians of their day. With this motivation, the modern theory of probability began to be developed.

Motivated by problems associated with games of chance, the theory of probability first was developed under the assumption of **equal likelihood**, expressed in Theorem 2.2. Under this assumption one only had to count the number of "successful" outcomes and divided by the total number of "possible" outcomes to arrive at the probability of an event.

The assumption of equal likelihood fails when we attempt, for example, to find the probability that a trifecta at the race track will pay off. Here, the different horses have different probabilities of winning, and we are forced to rely on a different method of evaluating probabilities. It is common to take into account the various horses' records in previous races, calculating each horse's probability of winning by dividing its number of wins by the number of starts. This idea gives rise to the **frequency interpretation** of probabilities, which interprets the probability of an event to be the proportion of times the event has occurred in a long series of repeated experiments. (This interpretation was first mentioned on page 24.) Application of the frequency interpretation requires a well-documented history of the outcomes of an event over a large number of experimental trials. In the absence of such a history, a series of experiments can be planned and their results observed. For example, the probability that a lot of manufactured items will have at most three defectives is estimated to be 0.90 if, in 90 percent of many previous lots *produced to the same specifications by the same process*, the number of defectives was three or less.

A more recently employed method of calculating probabilities is called the **subjective method**. Here, a personal, or subjective assessment is made of the probability of an event which is difficult or impossible to estimate in any other way. For example, the probability that the major stock market indexes will go up in a given future period of time cannot be estimated very well using the frequency interpretation because economic and world conditions rarely replicate themselves very closely. As another example, juries use this method when determining the guilt or innocence of a defendant "beyond a reasonable doubt." Subjective probabilities should be used only when all other methods fail, and then only with a high level of skepticism.

An important application of probability theory relates to the theory of **reliability**. The reliability of a component or system can be defined as follows.

> **Definition 2.4.** The **reliability** of a product is the probability that it will function within specified limits for a specified period of time under specified environmental conditions.

Thus, the reliability of a "standard equipment" automobile tire is close to 1 for 10,000 miles of operation on a passenger car traveling within the speed limits on paved roads, but it is close to zero for even short distances at the Indianapolis "500."

The reliability of a system of components can be calculated from the reliabilities of the individual components if the system consists entirely of components connected in series, or in parallel, or both. A **series system** is one in which all components are so interrelated that the entire system will fail if any one (or more) of its components fails. A **parallel system** will fail only if all its components fail. An example of a series system is a string of Christmas lights connected electrically "in series." If one bulb fails, the entire string will fail to light. Parallel systems are sometimes called "redundant" systems. For example, if the hydraulic system on a commercial aircraft that lowers the landing wheels fails, they may be lowered manually with a crank.

We shall assume that the components connected in a series system are independent, that is, the performance of one part does not effect the reliability of the others. Under this assumption, the reliability of a parallel system is given by an extension of Definition 2.2. Thus, we have

> **Theorem 2.14.** The **reliability of a series system** consisting of n independent components is given by
>
> $$R_s = \prod_{i=1}^{n} R_i$$
>
> where R_i is the reliability of the ith component.

Proof. The proof follows immediately by iterating in Definition 2.2. □

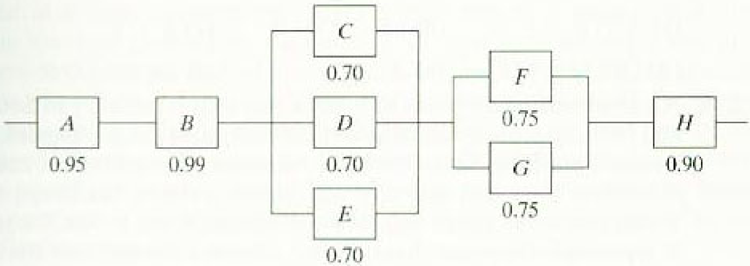

FIGURE 2.12: Combination of series and parallel systems.

Theorem 2.14 vividly demonstrates the effect of increased complexity on reliability. For example, if a series system has 5 components, each with a reliability of 0.970, the reliability of the entire system is only $(0.970)^5 = 0.859$. If the system complexity were increased so it now has 10 such components, the reliability would be reduced to $(0.970)^{10} = 0.738$.

One way to improve the reliability of a series system is to introduce parallel redundancy by replacing some or all of its components by several components connected in parallel. If a system consists of n independent components connected in parallel, it will fail to function only if all components fail. Thus, for the ith component, the probability of failure is $F_i = 1 - R_i$, called the "unreliability" of the component. Again applying Definition 2.2, we obtain

Theorem 2.15. The **reliability of a parallel system** consisting of n independent components is given by

$$R_p = 1 - \prod_{i=1}^{n}(1 - R_i)$$

Proof. The proof of this theorem is identical to that of Theorem 2.14, with $(1 - R_i)$ replacing R_i. □

EXAMPLE 2.29

Consider the system diagramed in Figure 2.12, which consists of eight components having the reliabilities shown in the figure. Find the reliability of the system.

Solution The parallel subsystem C, D, E can be replaced by an equivalent component, C' having the reliability $1 - (1 - 0.70)^3 = 0.973$. Likewise, F, G can be replaced by F' having the reliability $1 - (1 - 0.75)^2 = 0.9375$. Thus, the system is reduced to the parallel system A, B, C', F', H, having the reliability $(0.95)(0.99)(0.973)(0.9375)(0.90) = 0.772$. ■

APPLIED EXERCISES SECS. 2.1–2.3

2.35. If $S = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$, $A = \{1, 3, 5, 7\}$, $B = \{6, 7, 8, 9\}$, $C = \{2, 4, 8\}$, and $D = \{1, 5, 9\}$, list the elements of the subsets of S corresponding to the following events:

(a) $A' \cap B$; **(b)** $(A' \cap B) \cap C$; **(c)** $B' \cup C$;

(d) $(B' \cup C) \cap D$; **(e)** $A' \cap C$; **(f)** $(A' \cap C) \cap D$.

2.36. An electronics firm plans to build a research laboratory in Southern California, and its management has to decide between sites in Los Angeles, San Diego, Long Beach, Pasadena, Santa Barbara, Anaheim, Santa Monica, and Westwood. If A represents the event that they will choose a site in San Diego or Santa Barbara, B represents the event that they will choose a site in San Diego or Long Beach, C represents the event that they will choose a site in Santa Barbara or Anaheim, and D represents the event that they will choose a site in Los Angeles or Santa Barbara, list the elements of each of the following subsets of the sample space, which consists of the eight site selections:

(a) A'; **(b)** D'; **(c)** $C \cap D$;

(d) $B \cap C$; **(e)** $B \cup C$; **(f)** $A \cup B$;

(g) $C \cup D$; **(h)** $(B \cup C)'$; **(i)** $B' \cap C'$.

2.37. Among the eight cars that a dealer has in his showroom, Car 1 is new, has air-conditioning, power steering, and bucket seats; Car 2 is one year old, has air-conditioning, but neither power steering nor bucket seats; Car 3 is two years old, has air-conditioning and power steering, but no bucket seats; Car 4 is three years old, has air-conditioning, but neither power steering nor bucket seats; Car 5 is new, has no air-conditioning, no power steering, and no bucket seats; Car 6 is one year old, has power steering, but neither air-conditioning nor bucket seats; Car 7 is two years old, has no air-conditioning, no power steering, and no bucket seats; and Car 8 is three years old, has no air-conditioning, but has power steering as well as bucket seats. If a customer buys one of these cars and the event that he chooses a new car, for example, is represented by the set {Car 1, Car 5}, indicate similarly the sets that represent the events that

(a) he chooses a car without air-conditioning;
(b) he chooses a car without power steering;
(c) he chooses a car with bucket seats;
(d) he chooses a car that is either two or three years old.

2.38. With reference to Exercise 2.37, state in words what kind of car the customer will choose, if his choice is given by

(a) the complement of the set of part (a);
(b) the union of the sets of parts (b) and (c);
(c) the intersection of the sets of parts (c) and (d);
(d) the intersection of parts (b) and (c) of this exercise.

2.39. If Ms. Brown buys one of the houses advertised for sale in a Seattle newspaper (on a given Sunday), T is the event that the house has three or more baths, U is the event that it has a fireplace, V is the event that it costs more than \$100,000, and W is the event that it is new, describe (in words) each of the following events:

(a) T'; **(b)** U'; **(c)** V';

(d) W'; **(e)** $T \cap U$; **(f)** $T \cap V$;

(g) $U' \cap V$; **(h)** $V \cup W$; **(i)** $V' \cup W$;

(j) $T \cup U$; **(k)** $T \cup V$; **(l)** $V \cap W$.

2.40. A resort hotel has two station wagons, which it uses to shuttle its guests to and from the airport. If the larger of the two station wagons can carry five passengers and the smaller can carry four passengers, the point (0, 3) represents the event that at a given moment the larger station wagon is empty while the smaller one has three passengers, the point (4, 2) represents the event that at the given moment the larger station wagon has four passengers while the smaller one has two passengers, ..., draw a figure showing the 30 points of the corresponding sample space. Also, if E stands for the event that at least one of the station wagons is empty, F stands for the event that together they carry two, four, or six passengers, and G stands for the event that each carries the same number of passengers, list the points of the sample space that correspond to each of the following events:

(a) E; **(b)** F; **(c)** G;

(d) $E \cup F$; **(e)** $E \cap F$; **(f)** $F \cup G$;

(g) $E \cup F'$; **(h)** $E \cap G'$; **(i)** $F' \cap E'$.

2.41. A coin is tossed once. Then, if it comes up heads, a die is thrown once; if the coin comes up tails, it is tossed twice more. Using the notation in which (H, 2), for example, denotes the event that the coin comes up heads and then the die comes up 2, and (T, T, T) denotes the event that the coin comes up tails three times in a row, list

(a) the 10 elements of the sample space S;

(b) the elements of S corresponding to event A that exactly one head occurs;

(c) the elements of S corresponding to event B that at least two tails occur or a number greater than 4 occurs.

2.42. An electronic game contains three components arranged in the series-parallel circuit shown in Figure 2.13. At any given time, each component may or may not be operative, and the game will operate only if there is a continuous circuit from P to Q. Let A be the event that the game will operate; let B be the event that the game will operate though component x is not operative; and let C be the event that the game will operate though component y is not operative. Using the notation in which (0, 0, 1), for example, denotes that component z is operative but components x and y are not,

(a) list the elements of the sample space S and also the elements of S corresponding to events A, B, and C;

(b) determine which pairs of events, A and B, A and C, or B and C, are mutually exclusive.

2.43. An experiment consists of rolling a die until a 3 appears. Describe the sample space and determine

(a) how many elements of the sample space correspond to the event that the 3 appears on the kth roll of the die;

(b) how many elements of the sample space correspond to the event that the 3 appears not later than the kth roll of the die.

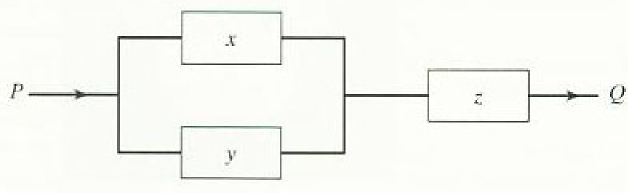

FIGURE 2.13: Diagram for Exercise 2.42.

2.44. If $S = \{x|0 < x < 10\}$, $M = \{x|3 < x \leqq 8\}$, and $N = \{x|5 < x < 10\}$, find

(a) $M \cup N$; **(b)** $M \cap N$;

(c) $M \cap N'$; **(d)** $M' \cup N$.

2.45. Express symbolically the sample space S that consists of all the points (x, y) on or in the circle of radius 3 centered at the point $(2, -3)$.

2.46. In Figure 2.14, L is the event that a driver has liability insurance and C is the event that she has collision insurance. Express in words what events are represented by regions 1, 2, 3, and 4.

2.47. With reference to Exercise 2.46 and Figure 2.14, what events are represented by

(a) regions 1 and 2 together;
(b) regions 2 and 4 together;
(c) regions 1, 2, and 3 together;
(d) regions 2, 3, and 4 together?

2.48. In Figure 2.15, E, T, and N are the events that a car brought to a garage needs an engine overhaul, transmission repairs, or new tires. Express in words the events represented by

(a) region 1;
(b) region 3;
(c) region 7;
(d) regions 1 and 4 together;
(e) regions 2 and 5 together;
(f) regions 3, 5, 6, and 8 together.

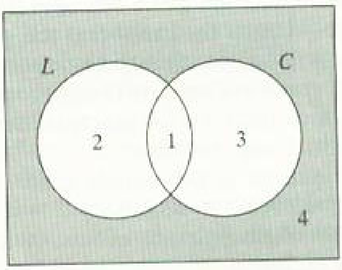

FIGURE 2.14: Venn diagram for Exercise 2.46.

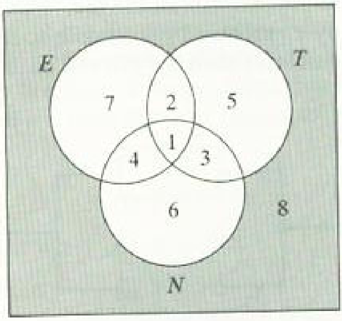

FIGURE 2.15: Venn diagram for Exercise 2.48.

2.49. With reference to Exercise 2.48 and Figure 2.15, list the region or combinations of regions representing the events that a car brought to the garage needs
(a) transmission repairs, but neither an engine overhaul nor new tires;
(b) an engine overhaul and transmission repairs;
(c) transmission repairs or new tires, but not an engine overhaul;
(d) new tires.

2.50. In a group of 200 college students, 138 are enrolled in a course in psychology, 115 are enrolled in a course in sociology, and 91 are enrolled in both. How many of these students are not enrolled in either course? (*Hint:* Draw a suitable Venn diagram and fill in the numbers associated with the various regions.)

2.51. A market research organization claims that, among 500 shoppers interviewed, 308 regularly buy Product X, 266 regularly buy Product Y, 103 regularly buy both, and 59 buy neither on a regular basis. Using a Venn diagram and filling in the number of shoppers associated with the various regions, check whether the results of this study should be questioned.

2.52. Among 120 visitors to Disneyland, 74 stayed for at least 3 hours, 86 spent at least \$20, 64 went on the Matterhorn ride, 60 stayed for at least 3 hours and spent at least \$20, 52 stayed for at least 3 hours and went on the Matterhorn ride, 54 spent at least \$20 and went on the Matterhorn ride, and 48 stayed for at least 3 hours, spent at least \$20, and went on the Matterhorn ride. Drawing a Venn diagram with three circles (like that of Figure 2.4) and filling in the numbers associated with the various regions, find how many of the 120 visitors to Disneyland
(a) stayed for at least 3 hours, spent at least \$20, but did not go on the Matterhorn ride;
(b) went on the Matterhorn ride, but stayed less than 3 hours and spent less than \$20;
(c) stayed less than 3 hours, spent at least \$20, but did not go on the Matterhorn ride.

SECS. 2.4–2.5

2.53. An experiment has five possible outcomes, A, B, C, D, and E, that are mutually exclusive. Check whether the following assignments of probabilities are permissible and explain your answers:
(a) $P(A) = 0.20$, $P(B) = 0.20$, $P(C) = 0.20$, $P(D) = 0.20$, and $P(E) = 0.20$;
(b) $P(A) = 0.21$, $P(B) = 0.26$, $P(C) = 0.58$, $P(D) = 0.01$, and $P(E) = 0.06$;
(c) $P(A) = 0.18$, $P(B) = 0.19$, $P(C) = 0.20$, $P(D) = 0.21$, and $P(E) = 0.22$;
(d) $P(A) = 0.10$, $P(B) = 0.30$, $P(C) = 0.10$, $P(D) = 0.60$, and $P(E) = -0.10$;
(e) $P(A) = 0.23$, $P(B) = 0.12$, $P(C) = 0.05$, $P(D) = 0.50$, and $P(E) = 0.08$.

2.54. If A and B are mutually exclusive, $P(A) = 0.37$, and $P(B) = 0.44$, find
(a) $P(A')$; **(b)** $P(B')$; **(c)** $P(A \cup B)$;
(d) $P(A \cap B)$; **(e)** $P(A \cap B')$; **(f)** $P(A' \cap B')$.

2.55. Explain why there must be a mistake in each of the following statements:
(a) The probability that Jean will pass the bar examination is 0.66 and the probability that she will not pass is -0.34.
(b) The probability that the home team will win an upcoming football game is 0.77, the probability that it will tie the game is 0.08, and the probability that it will win or tie the game is 0.95.
(c) The probabilities that a secretary will make 0, 1, 2, 3, 4, or 5 *or more* mistakes in typing a report are, respectively, 0.12, 0.25, 0.36, 0.14, 0.09, and 0.07.
(d) The probabilities that a bank will get 0, 1, 2, or 3 *or more* bad checks on any given day are, respectively, 0.08, 0.21, 0.29, and 0.40.

2.56. Suppose that each of the 30 points of the sample space of Exercise 2.40 is assigned the probability $\frac{1}{30}$. Find the probabilities that at a given moment

(a) at least one of the station wagons is empty;

(b) each of the two station wagons carries the same number of passengers;

(c) the larger station wagon carries more passengers than the smaller station wagon;

(d) together they carry at least six passengers.

2.57. The probabilities that the serviceability of a new X-ray machine will be rated very difficult, difficult, average, easy, or very easy are, respectively, 0.12, 0.17, 0.34, 0.29, and 0.08. Find the probabilities that the serviceability of the machine will be rated

(a) difficult or very difficult;

(b) neither very difficult nor very easy;

(c) average or worse;

(d) average or better.

2.58. A police department needs new tires for its patrol cars and the probabilities are 0.15, 0.24, 0.03, 0.28, 0.22, and 0.08 that it will buy Uniroyal tires, Goodyear tires, Michelin tires, General tires, Goodrich tires, or Armstrong tires. Find the probabilities that it will buy

(a) Goodyear or Goodrich tires;

(b) Uniroyal, Michelin, or Goodrich tires;

(c) Michelin or Armstrong tires;

(d) Uniroyal, Michelin, General, or Goodrich tires.

2.59. A hat contains twenty white slips of paper numbered from 1 through 20, ten red slips of paper numbered from 1 through 10, forty yellow slips of paper numbered from 1 through 40, and ten blue slips of paper numbered from 1 through 10. If these 80 slips of paper are thoroughly shuffled so that each slip has the same probability of being drawn, find the probabilities of drawing a slip of paper that is

(a) blue or white;

(b) numbered 1, 2, 3, 4, or 5;

(c) red or yellow and also numbered 1, 2, 3, or 4;

(d) numbered 5, 15, 25, or 35;

(e) white and numbered higher than 12 or yellow and numbered higher than 26.

2.60. Four candidates are seeking a vacancy on a school board. If A is twice as likely to be elected as B, and B and C are given about the same chance of being elected, while C is twice as likely to be elected as D, what are the probabilities that

(a) C will win;

(b) A will not win?

2.61. Two cards are randomly drawn from a deck of 52 playing cards. Find the probability that both cards will be greater than 3 and less than 8.

2.62. In a poker game, five cards are dealt at random from an ordinary deck of 52 playing cards. Find the probabilities of getting

(a) two pairs (any two distinct face values occurring exactly twice);

(b) four of a kind (four cards of equal face value).

2.63. In a game of Yahtzee, five balanced dice are rolled simultaneously. Find the probabilities of getting

(a) two pairs;

(b) three of a kind;

(c) a full house (three of a kind and a pair);

(d) four of a kind.

2.64. Among the 78 doctors on the staff of a hospital, 64 carry malpractice insurance, 36 are surgeons, and 34 of the surgeons carry malpractice insurance. If one of these doctors is chosen by lot to represent the hospital staff at an A.M.A. convention (that is, each doctor has a probability of $\frac{1}{78}$ of being selected), what is the probability that the one chosen is not a surgeon and does not carry malpractice insurance?

2.65. Explain on the basis of the various rules of Exercises 2.5 through 2.9 why there is a mistake in each of the following statements:

(a) The probability that it will rain is 0.67, and the probability that it will rain or snow is 0.55.

(b) The probability that a student will get a passing grade in English is 0.82, and the probability that she will get a passing grade in English and French is 0.86.

(c) The probability that a person visiting the San Diego Zoo will see the giraffes is 0.72, the probability that he will see the bears is 0.84, and the probability that he will see both is 0.52.

2.66. A line segment of length l is divided by a point selected at random within the segment. What is the probability that it will divide the line segment in a ratio greater than 1 : 2? What is the probability that it will divide the segment exactly in half?

2.67. A right triangle has the legs 3 and 4 units, respectively. Find the probability that a line segment, drawn at random parallel to the hypotenuse and contained entirely in the triangle, will divide the triangle so that the area between the line and the vertex opposite the hypotenuse will equal at least half the area of the triangle.

2.68. Given $P(A) = 0.59$, $P(B) = 0.30$, and $P(A \cap B) = 0.21$, find

(a) $P(A \cup B)$; **(b)** $P(A \cap B')$;

(c) $P(A' \cup B')$; **(d)** $P(A' \cap B')$.

2.69. For married couples living in a certain suburb, the probability that the husband will vote in a school board election is 0.21, the probability that the wife will vote in the election is 0.28, and the probability that they will both vote is 0.15. What is the probability that at least one of them will vote?

2.70. A biology professor has two graduate assistants helping her with her research. The probability that the older of the two assistants will be absent on any given day is 0.08, the probability that the younger of the two will be absent on any given day is 0.05, and the probability that they will both be absent on any given day is 0.02. Find the probabilities that

(a) either or both of the graduate assistants will be absent on any given day;

(b) at least one of the two graduate assistants will not be absent on any given day;

(c) only one of the two graduate assistants will be absent on any given day.

2.71. At Roanoke College it is known that $\frac{1}{3}$ of the students live off campus. It is also known that $\frac{5}{9}$ of the students are from within the state of Virginia and that $\frac{3}{4}$ of the students are from out of state or live on campus. What is the probability that a student selected at random from Roanoke College is from out of state and lives on campus?

2.72. Suppose that if a person visits Disneyland, the probability that he will go on the Jungle Cruise is 0.74, the probability that he will ride the Monorail is 0.70, the probability that he will go on the Matterhorn ride is 0.62, the probability that he will go on the Jungle Cruise and ride the Monorail is 0.52, the probability that he will go on the Jungle Cruise as well as the Matterhorn ride is 0.46, the probability that he will ride the Monorail and go on the Matterhorn ride is 0.44, and the probability that he will go on all three of these rides is 0.34. What is the probability that a person visiting Disneyland will go on at least one of these three rides?

2.73. Suppose that if a person travels to Europe for the first time, the probability that he will see London is 0.70, the probability that he will see Paris is 0.64, the probability that he will see Rome is 0.58, the probability that he will see Amsterdam is 0.58, the probability that he will see London and Paris is 0.45, the probability that he will see London and Rome is 0.42, the probability that he will see London and Amsterdam is 0.41, the probability that he will see Paris and Rome is 0.35, the probability that he will see Paris and Amsterdam is 0.39, the probability that he will see Rome and Amsterdam is 0.32, the probability that he will see London, Paris, and Rome is 0.23, the probability that he will see London, Paris, and Amsterdam is 0.26, the probability that he will see London, Rome, and Amsterdam is 0.21, the probability that he will see Paris, Rome, and Amsterdam is 0.20, and the probability that he will see all four of these cities is 0.12. What is the probability that a person traveling to Europe for the first time will see at least one of these four cities? (*Hint*: Use the formula of Exercise 2.13)

2.74. Use the formula of Exercise 2.15 to convert each of the following odds to probabilities:

(a) If three eggs are randomly chosen from a carton of 12 eggs of which three are cracked, the odds are 34 to 21 that at least one of them will be cracked.

(b) If a person has eight \$1 bills, five \$5 bills, and one \$20 bill, and randomly selects three of them, the odds are 11 to 2 that they will not all be \$1 bills.

(c) If we arbitrarily arrange the letters in the word "nest," the odds are 5 to 1 that we will not get a meaningful word in the English language.

2.75. Use the definition of "odds" given in Exercise 2.15 to convert each of the following probabilities to odds:

(a) The probability that the last digit of a car's license plate is a 2, 3, 4, 5, 6, or 7 is $\frac{6}{10}$.

(b) The probability of getting at least two heads in four flips of a balanced coin is $\frac{11}{16}$.

(c) The probability of rolling "7 or 11" with a pair of balanced dice is $\frac{2}{9}$.

SECS. 2.6–2.8

2.76. There are 90 applicants for a job with the news department of a television station. Some of them are college graduates and some are not, some of them have at least three years' experience and some have not, with the exact breakdown being

	College graduates	*Not college graduates*
At least three years' experience	18	9
Less than three years' experience	36	27

If the order in which the applicants are interviewed by the station manager is random, G is the event that the first applicant interviewed is a college graduate, and T is the event that the first applicant interviewed has at least three years' experience, determine each of the following probabilities directly from the entries and the row and column totals of the table:

(a) $P(G)$; **(b)** $P(T')$; **(c)** $P(G \cap T)$;

(d) $P(G' \cap T')$; **(e)** $P(T|G)$; **(f)** $P(G'|T')$.

2.77. Use the results of Exercise 2.76 to verify that

(a) $P(T|G) = \frac{P(G \cap T)}{P(G)}$;

(b) $P(G'|T') = \frac{P(G' \cap T')}{P(T')}$.

2.78. With reference to Exercise 2.64, what is the probability that the doctor chosen to represent the hospital staff at the convention carries malpractice insurance given that he or she is a surgeon?

2.79. With reference to Exercise 2.69, what is the probability that a husband will vote in the election given that his wife is going to vote?

2.80. With reference to Exercise 2.71, what is the probability that one of the students will be living on campus given that he or she is from out of state?

2.81. A bin contains 100 balls, of which 25 are red, 40 are white, and 35 are black. If two balls are selected from the bin without replacement, what is the probability that one will be red and one will be white?

2.82. If subjective probabilities are determined by the method suggested in Exercise 2.16, the third postulate of probability may not be satisfied. However, proponents of the subjective probability concept usually impose this postulate as a **consistency criterion**; in other words, they regard subjective probabilities that do not satisfy the postulate as inconsistent.

(a) A high school principal feels that the odds are 7 to 5 against her getting a $1,000 raise and 11 to 1 against her getting a $2,000 raise. Furthermore, she feels that it is an even-money bet that she will get one of these raises or the other. Discuss the consistency of the corresponding subjective probabilities.

(b) Asked about his political future, a party official replies that the odds are 2 to 1 that he will not run for the House of Representatives and 4 to 1 that he will not run for the Senate. Furthermore, he feels that the odds are 7 to 5 that he will run for one or the other. Are the corresponding probabilities consistent?

2.83. There are two Porsches in a road race in Italy, and a reporter feels that the odds against their winning are 3 to 1 and 5 to 3. To be consistent (see Exercise 2.82), what odds should the reporter assign to the event that either car will win?

2.84. If we let x = the number of spots facing up when a pair of dice is cast, then we can use the sample space S_2 of Example 2.2 to describe the outcomes of the experiment.

(a) Find the probability of each outcome in S_2.

(b) Verify that the sum of these probabilities is 1.

2.85. Using a computer program that can generate random integers on the interval (0, 9) with equal probabilities, generate 1,000 such integers and use the frequency interpretation to estimate the probability that such a randomly chosen integer will have a value less than 1.

2.86. Using the method of Exercise 2.85, generate a second set of 1,000 random integers on (0, 9). Estimate the probability that A: an integer selected at random from the first set will be less than 1 or B: an integer selected at random from the second set will be less than 1

(a) using the frequency interpretation of probabilities;

(b) using Theorem 2.7 and $P(A \cap B) = 0.25$.

2.87. It is felt that the probabilities are 0.20, 0.40, 0.30, and 0.10 that the basketball teams of four universities, T, U, V, and W, will win their conference championship. If university U is placed on probation and declared ineligible for the championship, what is the probability that university T will win the conference championship?

2.88. With reference to Exercise 2.72, find the probabilities that a person who visits Disneyland will

(a) ride the Monorail given that he will go on the Jungle Cruise;

(b) go on the Matterhorn ride given that he will go on the Jungle Cruise and ride the Monorail;

(c) not go on the Jungle Cruise given that he will ride the Monorail and/or go on the Matterhorn ride;

(d) go on the Matterhorn ride and the Jungle Cruise given that he will not ride the Monorail.

(*Hint*: Draw a Venn diagram and fill in the probabilities associated with the various regions.)

2.89. The probability of surviving a certain transplant operation is 0.55. If a patient survives the operation, the probability that his or her body will reject the transplant within a month is 0.20. What is the probability of surviving both of these critical stages?

2.90. Crates of eggs are inspected for blood clots by randomly removing three eggs in succession and examining their contents. If all three eggs are good, the crate is shipped; otherwise it is rejected. What is the probability that a crate will be shipped if it contains 120 eggs, of which 10 have blood clots?

2.91. Suppose that in Vancouver, B.C., the probability that a rainy fall day is followed by a rainy day is 0.80 and the probability that a sunny fall day is followed by a rainy day is 0.60. Find the probabilities that a rainy fall day is followed by

(a) a rainy day, a sunny day, and another rainy day;

(b) two sunny days and then a rainy day;

(c) two rainy days and then two sunny days;

(d) rain two days later.

[*Hint*: In part (c) use the formula of Exercise 2.19.]

2.92. Use the formula of Exercise 2.19 to find the probability of randomly choosing (without replacement) four healthy guinea pigs from a cage containing 20 guinea pigs, of which 15 are healthy and 5 are diseased.

2.93. A balanced die is tossed twice. If A is the event that an even number comes up on the first toss, B is the event that an even number comes up on the second toss, and C is the event that both tosses result in the same number, are the events A, B, and C

(a) pairwise independent;

(b) independent?

2.94. A sharpshooter hits a target with probability 0.75. Assuming independence, find the probabilities of getting

(a) a hit followed by two misses;

(b) two hits and a miss in any order.

2.95. A coin is loaded so that the probabilities of heads and tails are 0.52 and 0.48, respectively. If the coin is tossed three times, what are the probabilities of getting

(a) all heads;

(b) two tails and a head in that order?

2.96. A shipment of 1,000 parts contains 1 percent defective parts. Find the probability that

(a) the first four parts chosen arbitrarily for inspection are nondefective;

(b) the first defective part found will be on the fourth inspection.

2.97. Medical records show that one out of 10 persons in a certain town has a thyroid deficiency. If 12 persons in this town are randomly chosen and tested, what is the probability that at least one of them will have a thyroid deficiency?

2.98. If five of a company's 10 delivery trucks do not meet emission standards and three of them are chosen for inspection, what is the probability that none of the trucks chosen will meet emission standards?

2.99. If a person randomly picks four of the 15 gold coins a dealer has in stock, and six of the coins are counterfeits, what is the probability that the coins picked will all be counterfeits?

2.100. A department store that bills its charge-account customers once a month has found that if a customer pays promptly one month, the probability is 0.90 that he or she will also pay promptly the next month; however, if a customer does not pay promptly one month, the probability that he or she will pay promptly the next month is only 0.40.

(a) What is the probability that a customer who pays promptly one month will also pay promptly the next three months?

(b) What is the probability that a customer who does not pay promptly one month will also not pay promptly the next two months and then make a prompt payment the month after that?

2.101. With reference to Figure 2.16, verify that events A, B, C, and D are independent. Note that the region representing A consists of two circles, and so do the regions representing B and C.

2.102. At an electronics plant, it is known from past experience that the probability is 0.84 that a new worker who has attended the company's training program will meet the production quota, and that the corresponding probability is 0.49 for a new worker who has not attended the company's training program. If 70 percent of all new workers attend the training program, what is the probability that a new worker will meet the production quota?

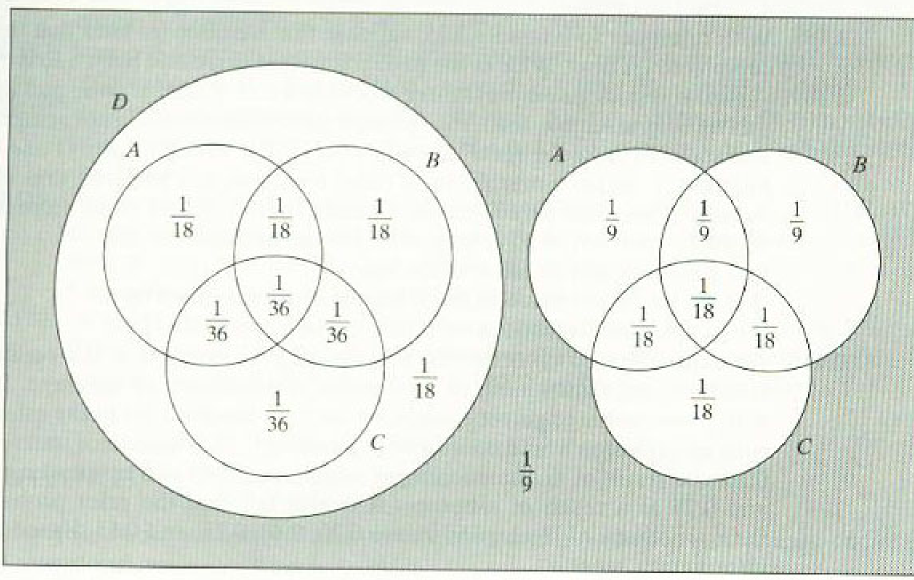

FIGURE 2.16: Diagram for Exercise 2.101.

2.103. In a T-maze, a rat is given food if it turns left and an electric shock if it turns right. On the first trial there is a 50–50 chance that a rat will turn either way; then, if it receives food on the first trial, the probability is 0.68 that it will turn left on the next trial, and if it receives a shock on the first trial, the probability is 0.84 that it will turn left on the next trial. What is the probability that a rat will turn left on the second trial?

2.104. It is known from experience that in a certain industry 60 percent of all labor–management disputes are over wages, 15 percent are over working conditions, and 25 percent are over fringe issues. Also, 45 percent of the disputes over wages are resolved without strikes, 70 percent of the disputes over working conditions are resolved without strikes, and 40 percent of the disputes over fringe issues are resolved without strikes. What is the probability that a labor–management dispute in this industry will be resolved without a strike?

2.105. With reference to Exercise 2.104, what is the probability that if a labor–management dispute in this industry is resolved without a strike, it was over wages?

2.106. The probability that a one-car accident is due to faulty brakes is 0.04, the probability that a one-car accident is correctly attributed to faulty brakes is 0.82, and the probability that a one-car accident is incorrectly attributed to faulty brakes is 0.03. What is the probability that

(a) a one-car accident will be attributed to faulty brakes;

(b) a one-car accident attributed to faulty brakes was actually due to faulty brakes?

2.107. In a certain community, 8 percent of all adults over 50 have diabetes. If a health service in this community correctly diagnoses 95 percent of all persons with diabetes as having the disease and incorrectly diagnoses 2 percent of all persons without diabetes as having the disease, find the probabilities that

(a) the community health service will diagnose an adult over 50 as having diabetes;

(b) a person over 50 diagnosed by the health service as having diabetes actually has the disease.

2.108. With reference to Example 2.25, suppose that we discover later that the job was completed on time. What is the probability that there had been a strike?

2.109. A mail-order house employs three stock clerks, U, V, and W, who pull items from shelves and assemble them for subsequent verification and packaging. U makes a mistake in an order (gets a wrong item or the wrong quantity) one time in a hundred, V makes a mistake in an order five times in a hundred, and W makes a mistake in an order three times in a hundred. If U, V, and W fill, respectively, 30, 40, and 30 percent of all orders, what are the probabilities that

(a) a mistake will be made in an order;

(b) if a mistake is made in an order, the order was filled by U;

(c) if a mistake is made in an order, the order was filled by V?

2.110. An explosion at a construction site could have occurred as the result of static electricity, malfunctioning of equipment, carelessness, or sabotage. Interviews with construction engineers analyzing the risks involved led to the estimates that such an explosion would occur with probability 0.25 as a result of static electricity, 0.20 as a result of malfunctioning of equipment, 0.40 as a result of carelessness, and 0.75 as a result of sabotage. It is also felt that the prior probabilities of the four causes of the explosion are 0.20, 0.40, 0.25, and 0.15. Based on all this information, what is

(a) the most likely cause of the explosion;

(b) the least likely cause of the explosion?

2.111. An art dealer receives a shipment of five old paintings from abroad, and, on the basis of past experience, she feels that the probabilities are, respectively, 0.76, 0.09, 0.02, 0.01, 0.02, and 0.10 that 0, 1, 2, 3, 4, or all 5 of them are forgeries. Since the cost of authentication is fairly high, she decides to select one of the five paintings at random and send it away for authentication. If it turns out that this painting is a forgery, what probability should she now assign to the possibility that all the other paintings are also forgeries?

2.112. To get answers to sensitive questions, we sometimes use a method called the **randomized response technique.** Suppose, for instance, that we want to determine what percentage of the students at a large university smoke marijuana. We construct 20 flash cards, write "I smoke marijuana at least once a week" on 12 of the cards, where 12 is an arbitrary choice, and "I do not smoke marijuana at least once a week" on the others. Then, we let each student (in the sample interviewed) select one of the 20 cards at random and respond "yes" or "no" without divulging the question.

(a) Establish a relationship between $P(Y)$, the probability that a student will give a "yes" response, and $P(M)$, the probability that a student randomly selected at that university smokes marijuana at least once a week.

(b) If 106 of 250 students answered "yes" under these conditions, use the result of part (a) and $\frac{106}{250}$ as an estimate of $P(Y)$ to estimate $P(M)$.

SEC. 2.9

2.113. Find the reliability of a series systems having five components with reliabilities 0.995, 0.990, 0.992, 0.995, 0.998, respectively.

2.114. A series system consists of three components, each having the reliability 0.95, and three components, each having the reliability 0.99. Find the reliability of the system.

2.115. What must be the reliability of each component in a series system consisting of six components that must have a system reliability of 0.95?

2.116. Referring to Exercise 2.115, suppose now that there are ten components, and the system reliability must be 0.90.

2.117. Suppose a system consists of four components, connected in parallel, having the reliabilities 0.8, 0.7, 0.7, and 0.65, respectively. Find the system reliability.

2.118. Referring to Exercise 2.117, suppose now that the system has five components with reliabilities 0.85, 0.80, 0.65, 0.60, and 0.70, respectively. Find the system reliability.

2.119. A system consists of two components having the reliabilities 0.95 and 0.90, connected in series to two parallel subsystems, the first containing four components, each having the reliability 0.60 and the second containing two components, each having the reliability 0.75. Find the system reliability.

2.120. A series system consists of two components having the reliabilities 0.98 and 0.99 connected to a parallel subsystem containing five components having the reliabilities 0.75, 0.60, 0.65, 0.70, and 0.60. Find the system reliability.

REFERENCES

Among the numerous textbooks on probability theory published in recent years, one of the most popular is

FELLER, W., *An Introduction to Probability Theory and Its Applications*, Vol. I, 3rd ed. New York: John Wiley & Sons, Inc., 1968.

More elementary treatments may be found in

BARR, D. R., and ZEHNA, P. W., *Probability: Modeling Uncertainty*. Reading, Mass.: Addison-Wesley Publishing Company, Inc., 1983.

DRAPER, N. R., and LAWRENCE, W. E., *Probability: An Introductory Course*. Chicago: Markam Publishing Company, 1970.

FREUND, J. E., *Introduction to Probability*. New York: Dover Publications, Inc., 1993 Reprint.

GOLDBERG, S., *Probability—An Introduction*. Mineola, N.Y.: Dover Publications, Inc. (republication of 1960 edition).

HODGES, J. L., and LEHMANN, E. L., *Elements of Finite Probability*. San Francisco: Holden Day, Inc., 1965.

NOSAL, M., *Basic Probability and Applications*. Philadelphia: W. B. Saunders Company, 1977.

More advanced treatments are given in many texts, for instance,

HOEL, P., PORT, S. C., and STONE, C. J., *Introduction to Probability Theory*. Boston: Houghton Mifflin Company, 1971.

KHAZANIE, R., *Basic Probability Theory and Applications*. Pacific Palisades, Calif.: Goodyear Publishing Company, Inc., 1976.

PARZEN, E., *Modern Probability Theory and Its Applications*. New York: John Wiley & Sons, Inc., 1960.

ROSS, S., *A First Course in Probability*, 3rd ed. New York: Macmillan Publishing Company, 1988.

SOLOMON, F., *Probability and Stochastic Processes*. Upper Saddle River, N.J.: Prentice Hall, 1987.

For more on reliability and related topics, see

JOHNSON, R.A., *Miller and Freund's Probability and Statistics for Engineers*, Upper Saddle River, N.J.: Prentice Hall, 2000.

MILLER, I. and MILLER, M., *Statistical Methods for Quality with Applications to Engineering and Management*, Upper Saddle Reiver, N.J.: Prentice Hall, 1995.