

Chapter 2

Probability

2.1 Sample Space

In the study of statistics, we are concerned basically with the presentation and interpretation of **chance outcomes** that occur in a planned study or scientific investigation. For example, we may record the number of accidents that occur monthly at the intersection of Driftwood Lane and Royal Oak Drive, hoping to justify the installation of a traffic light; we might classify items coming off an assembly line as "defective" or "nondefective"; or we may be interested in the volume of gas released in a chemical reaction when the concentration of an acid is varied. Hence, the statistician is often dealing with either numerical data, representing counts or measurements, or **categorical data**, which can be classified according to some criterion.

We shall refer to any recording of information, whether it be numerical or categorical, as an **observation**. Thus, the numbers 2, 0, 1, and 2, representing the number of accidents that occurred for each month from January through April during the past year at the intersection of Driftwood Lane and Royal Oak Drive, constitute a set of observations. Similarly, the categorical data N , D , N , N , and D , representing the items found to be defective or nondefective when five items are inspected, are recorded as observations.

Statisticians use the word **experiment** to describe any process that generates a set of data. A simple example of a statistical experiment is the tossing of a coin. In this experiment, there are only two possible outcomes, heads or tails. Another experiment might be the launching of a missile and observing of its velocity at specified times. The opinions of voters concerning a new sales tax can also be considered as observations of an experiment. We are particularly interested in the observations obtained by repeating the experiment several times. In most cases, the outcomes will depend on chance and, therefore, cannot be predicted with certainty. If a chemist runs an analysis several times under the same conditions, he or she will obtain different measurements, indicating an element of chance in the experimental procedure. Even when a coin is tossed repeatedly, we cannot be certain that a given toss will result in a head. However, we know the entire set of possibilities for each toss.

Given the discussion in Section 1.7, we should deal with the breadth of the term **experiment**. Three types of statistical studies were reviewed, and several examples were given of each. In each of the three cases, *designed experiments*, *observational studies*, and *retrospective studies*, the end result was a set of *data* that of course is

subject to uncertainty. Though only one of these has the word *experiment* in its description, the process of generating the data or the process of observing the data is part of an experiment. The corrosion study discussed in Section 1.2 certainly involves an experiment, with measures of corrosion representing the data. The example given in Section 1.7 in which blood cholesterol and sodium were observed on a group of individuals represented an observational study (as opposed to a *designed* experiment), and yet the process generated data and the outcome is subject to uncertainty. Thus, it is an experiment. A third example in Section 1.7 represented a retrospective study in which historical data on monthly electric power consumption and average monthly ambient temperature were observed. Even though the data may have been in the files for decades, the process is still referred to as an experiment.

Definition 2.1: The set of all possible outcomes of a statistical experiment is called the **sample space** and is represented by the symbol S .

Each outcome in a sample space is called an **element** or a **member** of the sample space, or simply a **sample point**. If the sample space has a finite number of elements, we may *list* the members separated by commas and enclosed in braces. Thus, the sample space S , of possible outcomes when a coin is flipped, may be written

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

where H and T correspond to heads and tails, respectively.

Example 2.1: Consider the experiment of tossing a die. If we are interested in the number that shows on the top face, the sample space is

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

If we are interested only in whether the number is even or odd, the sample space is simply

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

Example 2.1 illustrates the fact that more than one sample space can be used to describe the outcomes of an experiment. In this case, S_1 provides more information than S_2 . If we know which element in S_1 occurs, we can tell which outcome in S_2 occurs; however, a knowledge of what happens in S_2 is of little help in determining which element in S_1 occurs. In general, it is desirable to use the sample space that gives the most information concerning the outcomes of the experiment. In some experiments, it is helpful to list the elements of the sample space systematically by means of a **tree diagram**.

Example 2.2: An experiment consists of flipping a coin and then flipping it a second time if a head occurs. If a tail occurs on the first flip, then a die is tossed once. To list the elements of the sample space providing the most information, we construct the tree diagram of Figure 2.1. The various paths along the branches of the tree give the distinct sample points. Starting with the top left branch and moving to the right along the first path, we get the sample point HH , indicating the possibility that heads occurs on two successive flips of the coin. Likewise, the sample point $T3$ indicates the possibility that the coin will show a tail followed by a 3 on the toss of the die. By proceeding along all paths, we see that the sample space is

$$S = \{HH, HT, T1, T2, T3, T4, T5, T6\}.$$

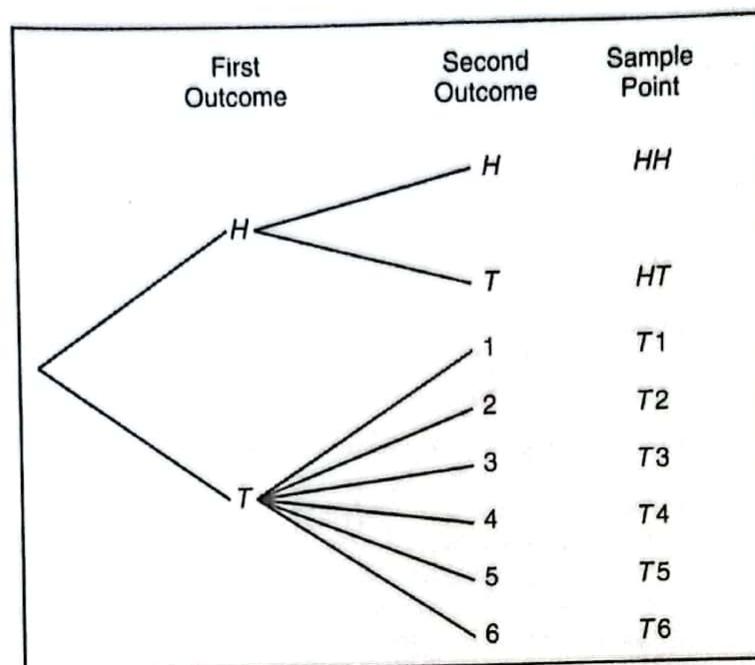


Figure 2.1: Tree diagram for Example 2.2.

Many of the concepts in this chapter are best illustrated with examples involving the use of dice and cards. These are particularly important applications to use early in the learning process, to facilitate the flow of these new concepts into scientific and engineering examples such as the following.

Example 2.3: Suppose that three items are selected at random from a manufacturing process. Each item is inspected and classified defective, D , or nondefective, N . To list the elements of the sample space providing the most information, we construct the tree diagram of Figure 2.2. Now, the various paths along the branches of the tree give the distinct sample points. Starting with the first path, we get the sample point DDD , indicating the possibility that all three items inspected are defective. As we proceed along the other paths, we see that the sample space is

$$S = \{DDD, DDN, DND, DNN, NDD, NDN, NND, NNN\}.$$

Sample spaces with a large or infinite number of sample points are best described by a **statement or rule method**. For example, if the possible outcomes of an experiment are the set of cities in the world with a population over 1 million, our sample space is written

$$S = \{x \mid x \text{ is a city with a population over 1 million}\},$$

which reads “ S is the set of all x such that x is a city with a population over 1 million.” The vertical bar is read “such that.” Similarly, if S is the set of all points (x, y) on the boundary or the interior of a circle of radius 2 with center at the origin, we write the rule

$$S = \{(x, y) \mid x^2 + y^2 \leq 4\}.$$

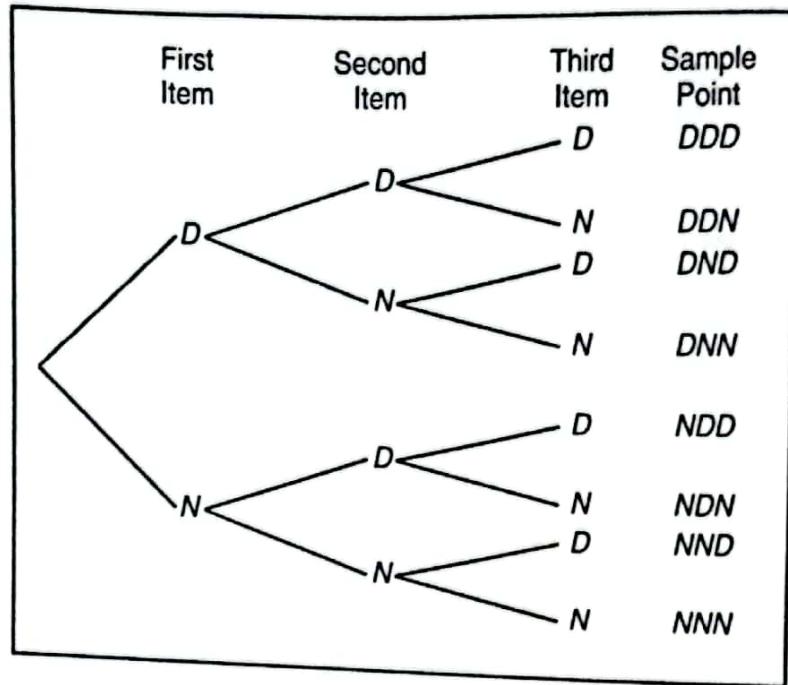


Figure 2.2: Tree diagram for Example 2.3.

Whether we describe the sample space by the rule method or by listing the elements will depend on the specific problem at hand. The rule method has practical advantages, particularly for many experiments where listing becomes a tedious chore.

Consider the situation of Example 2.3 in which items from a manufacturing process are either D , defective, or N , nondefective. There are many important statistical procedures called sampling plans that determine whether or not a "lot" of items is considered satisfactory. One such plan involves sampling until k defectives are observed. Suppose the experiment is to sample items randomly until one defective item is observed. The sample space for this case is

$$S = \{D, ND, NND, NNND, \dots\}.$$

2.2 Events

For any given experiment, we may be interested in the occurrence of certain events rather than in the occurrence of a specific element in the sample space. For instance, we may be interested in the event A that the outcome when a die is tossed is divisible by 3. This will occur if the outcome is an element of the subset $A = \{3, 6\}$ of the sample space S_1 in Example 2.1. As a further illustration, we may be interested in the event B that the number of defectives is greater than 1 in Example 2.3. This will occur if the outcome is an element of the subset

$$B = \{DDN, DND, NDD, DDD\}$$

of the sample space S .

To each event we assign a collection of sample points, which constitute a subset of the sample space. That subset represents all of the elements for which the event is true.

Definition 2.2: An event is a subset of a sample space.

Example 2.4: Given the sample space $S = \{t \mid t \geq 0\}$, where t is the life in years of a certain electronic component, then the event A that the component fails before the end of the fifth year is the subset $A = \{t \mid 0 \leq t < 5\}$.

It is conceivable that an event may be a subset that includes the entire sample space S or a subset of S called the null set and denoted by the symbol \emptyset , which contains no elements at all. For instance, if we let A be the event of detecting a microscopic organism by the naked eye in a biological experiment, then $A = \emptyset$. Also, if

$$B = \{x \mid x \text{ is an even factor of } 7\},$$

then B must be the null set, since the only possible factors of 7 are the odd numbers 1 and 7.

Consider an experiment where the smoking habits of the employees of a manufacturing firm are recorded. A possible sample space might classify an individual as a nonsmoker, a light smoker, a moderate smoker, or a heavy smoker. Let the subset of smokers be some event. Then all the nonsmokers correspond to a different event, also a subset of S , which is called the complement of the set of smokers.

Definition 2.3: The complement of an event A with respect to S is the subset of all elements of S that are not in A . We denote the complement of A by the symbol A' .

Example 2.5: Let R be the event that a red card is selected from an ordinary deck of 52 playing cards, and let S be the entire deck. Then R' is the event that the card selected from the deck is not a red card but a black card.

Example 2.6: Consider the sample space

$$S = \{\text{book, cell phone, mp3, paper, stationery, laptop}\}.$$

Let $A = \{\text{book, stationery, laptop, paper}\}$. Then the complement of A is $A' = \{\text{cell phone, mp3}\}$.

We now consider certain operations with events that will result in the formation of new events. These new events will be subsets of the same sample space as the given events. Suppose that A and B are two events associated with an experiment. In other words, A and B are subsets of the same sample space S . For example, in the tossing of a die we might let A be the event that an even number occurs and B the event that a number greater than 3 shows. Then the subsets $A = \{2, 4, 6\}$ and $B = \{4, 5, 6\}$ are subsets of the same sample space

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

Note that both A and B will occur on a given toss if the outcome is an element of the subset $\{4, 6\}$, which is just the intersection of A and B .

Definition 2.4: The intersection of two events A and B , denoted by the symbol $A \cap B$, is the event containing all elements that are common to A and B .

Example 2.7: Let E be the event that a person selected at random in a classroom is majoring in engineering, and let F be the event that the person is female. Then $E \cap F$ is the event of all female engineering students in the classroom.

Example 2.8: Let $V = \{a, e, i, o, u\}$ and $C = \{l, r, s, t\}$; then it follows that $V \cap C = \emptyset$. That is, V and C have no elements in common and, therefore, cannot both simultaneously occur.

For certain statistical experiments it is by no means unusual to define two events, A and B , that cannot both occur simultaneously. The events A and B are then said to be **mutually exclusive**. Stated more formally, we have the following definition:

Definition 2.5: Two events A and B are **mutually exclusive**, or **disjoint**, if $A \cap B = \emptyset$, that is, if A and B have no elements in common.

Example 2.9: A cable television company offers programs on eight different channels, three of which are affiliated with ABC, two with NBC, and one with CBS. The other two are an educational channel and the ESPN sports channel. Suppose that a person subscribing to this service turns on a television set without first selecting the channel. Let A be the event that the program belongs to the NBC network and B the event that it belongs to the CBS network. Since a television program cannot belong to more than one network, the events A and B have no programs in common. Therefore, the intersection $A \cap B$ contains no programs, and consequently the events A and B are mutually exclusive.

Often one is interested in the occurrence of at least one of two events associated with an experiment. Thus, in the die-tossing experiment, if

$$A = \{2, 4, 6\} \text{ and } B = \{4, 5, 6\},$$

we might be interested in either A or B occurring or both A and B occurring. Such an event, called the **union** of A and B , will occur if the outcome is an element of the subset $\{2, 4, 5, 6\}$.

Definition 2.6: The **union** of the two events A and B , denoted by the symbol $A \cup B$, is the event containing all the elements that belong to A or B or both.

Example 2.10: Let $A = \{a, b, c\}$ and $B = \{b, c, d, e\}$; then $A \cup B = \{a, b, c, d, e\}$.

Example 2.11: Let P be the event that an employee selected at random from an oil drilling company smokes cigarettes. Let Q be the event that the employee selected drinks alcoholic beverages. Then the event $P \cup Q$ is the set of all employees who either drink or smoke or do both.

Example 2.12: If $M = \{x \mid 3 < x < 9\}$ and $N = \{y \mid 5 < y < 12\}$, then

$$M \cup N = \{z \mid 3 < z < 12\}.$$

The relationship between events and the corresponding sample space can be illustrated graphically by means of **Venn diagrams**. In a Venn diagram we let the sample space be a rectangle and represent events by circles drawn inside the rectangle. Thus, in Figure 2.3, we see that

$$\begin{aligned} A \cap B &= \text{regions 1 and 2,} \\ B \cap C &= \text{regions 1 and 3,} \end{aligned}$$

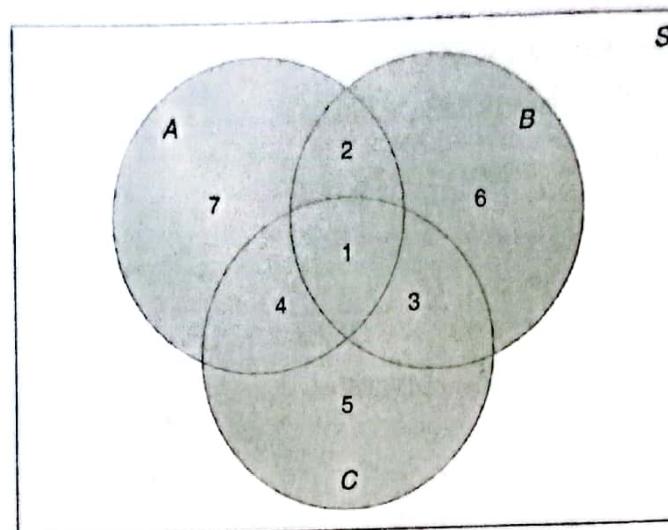


Figure 2.3: Events represented by various regions.

$$A \cup C = \text{regions } 1, 2, 3, 4, 5, \text{ and } 7,$$

$$B' \cap A = \text{regions } 4 \text{ and } 7,$$

$$A \cap B \cap C = \text{region } 1,$$

$$(A \cup B) \cap C' = \text{regions } 2, 6, \text{ and } 7,$$

and so forth.

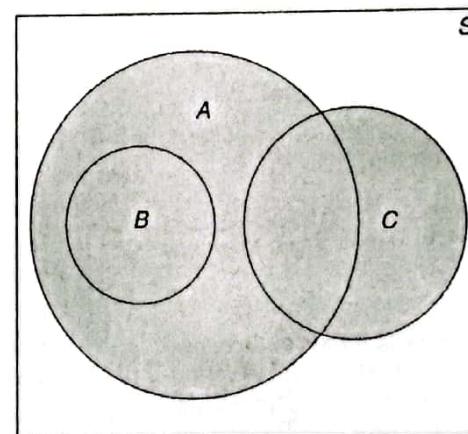


Figure 2.4: Events of the sample space S .

In Figure 2.4, we see that events A , B , and C are all subsets of the sample space S . It is also clear that event B is a subset of event A ; event $B \cap C$ has no elements and hence B and C are mutually exclusive; event $A \cap C$ has at least one element; and event $A \cup B = A$. Figure 2.4 might, therefore, depict a situation where we select a card at random from an ordinary deck of 52 playing cards and observe whether the following events occur:

A : the card is red,

B: the card is the jack, queen, or king of diamonds,

C: the card is an ace.

Clearly, the event $A \cap C$ consists of only the two red aces.

Several results that follow from the foregoing definitions, which may easily be verified by means of Venn diagrams, are as follows:

1. $A \cap \phi = \phi$.
2. $A \cup \phi = A$.
3. $A \cap A' = \phi$.
4. $A \cup A' = S$.
5. $S' = \phi$.

6. $\phi' = S$.
7. $(A')' = A$.
8. $(A \cap B)' = A' \cup B'$.
9. $(A \cup B)' = A' \cap B'$.

Exercises

2.1 List the elements of each of the following sample spaces:

- (a) the set of integers between 1 and 50 divisible by 8;
- (b) the set $S = \{x \mid x^2 + 4x - 5 = 0\}$;
- (c) the set of outcomes when a coin is tossed until a tail or three heads appear;
- (d) the set $S = \{x \mid x \text{ is a continent}\}$;
- (e) the set $S = \{x \mid 2x - 4 \geq 0 \text{ and } x < 1\}$.

2.2 Use the rule method to describe the sample space S consisting of all points in the first quadrant inside a circle of radius 3 with center at the origin.

2.3 Which of the following events are equal?

- (a) $A = \{1, 3\}$;
- (b) $B = \{x \mid x \text{ is a number on a die}\}$;
- (c) $C = \{x \mid x^2 - 4x + 3 = 0\}$;
- (d) $D = \{x \mid x \text{ is the number of heads when six coins are tossed}\}$.

2.4 An experiment involves tossing a pair of dice, one green and one red, and recording the numbers that come up. If x equals the outcome on the green die and y the outcome on the red die, describe the sample space S

- (a) by listing the elements (x, y) ;
- (b) by using the rule method.

2.5 An experiment consists of tossing a die and then flipping a coin once if the number on the die is even. If the number on the die is odd, the coin is flipped twice. Using the notation $4H$, for example, to denote the outcome that the die comes up 4 and then the coin comes up heads, and $3HT$ to denote the outcome that the die

comes up 3 followed by a head and then a tail on the coin, construct a tree diagram to show the 18 elements of the sample space S .

2.6 Two jurors are selected from 4 alternates to serve at a murder trial. Using the notation A_1A_3 , for example, to denote the simple event that alternates 1 and 3 are selected, list the 6 elements of the sample space S .

2.7 Four students are selected at random from a chemistry class and classified as male or female. List the elements of the sample space S_1 , using the letter M for male and F for female. Define a second sample space S_2 where the elements represent the number of females selected.

2.8 For the sample space of Exercise 2.4,

- (a) list the elements corresponding to the event A that the sum is greater than 8;
- (b) list the elements corresponding to the event B that a 2 occurs on either die;
- (c) list the elements corresponding to the event C that a number greater than 4 comes up on the green die;
- (d) list the elements corresponding to the event $A \cap C$;
- (e) list the elements corresponding to the event $A \cap B$;
- (f) list the elements corresponding to the event $B \cap C$;
- (g) construct a Venn diagram to illustrate the intersections and unions of the events A , B , and C .

2.9 For the sample space of Exercise 2.5,

- (a) list the elements corresponding to the event A that a number less than 3 occurs on the die;
- (b) list the elements corresponding to the event B that two tails occur;
- (c) list the elements corresponding to the event A' ;

- (d) list the elements corresponding to the event $A' \cap B$;
 (e) list the elements corresponding to the event $A \cup B$.

2.10 An engineering firm is hired to determine if certain waterways in Virginia are safe for fishing. Samples are taken from three rivers.

- (a) List the elements of a sample space S , using the letters F for safe to fish and N for not safe to fish.
 (b) List the elements of S corresponding to event E that at least two of the rivers are safe for fishing.
 (c) Define an event that has as its elements the points

$$\{FFF, NFF, FFN, NFN\}.$$

2.11 The resumés of two male applicants for a college teaching position in chemistry are placed in the same file as the resumés of two female applicants. Two positions become available, and the first, at the rank of assistant professor, is filled by selecting one of the four applicants at random. The second position, at the rank of instructor, is then filled by selecting at random one of the remaining three applicants. Using the notation M_2F_1 , for example, to denote the simple event that the first position is filled by the second male applicant and the second position is then filled by the first female applicant,

- (a) list the elements of a sample space S ;
 (b) list the elements of S corresponding to event A that the position of assistant professor is filled by a male applicant;
 (c) list the elements of S corresponding to event B that exactly one of the two positions is filled by a male applicant;
 (d) list the elements of S corresponding to event C that neither position is filled by a male applicant;
 (e) list the elements of S corresponding to the event $A \cap B$;
 (f) list the elements of S corresponding to the event $A \cup C$;
 (g) construct a Venn diagram to illustrate the intersections and unions of the events A , B , and C

2.12 Exercise and diet are being studied as possible substitutes for medication to lower blood pressure. Three groups of subjects will be used to study the effect of exercise. Group 1 is sedentary, while group 2 walks and group 3 swims for 1 hour a day. Half of each of the three exercise groups will be on a salt-free diet. An additional group of subjects will not exercise or restrict their salt, but will take the standard medication. Use Z for sedentary, W for walker, S for swimmer, Y for salt, N for no salt, M for medication, and F for medication free.

- (a) Show all of the elements of the sample space S .

- (b) Given that A is the set of nonmedicated subjects and B is the set of walkers, list the elements of $A \cup B$.

- (c) List the elements of $A \cap B$.

2.13 Construct a Venn diagram to illustrate the possible intersections and unions for the following events relative to the sample space consisting of all automobiles made in the United States.

$$F : \text{Four door}, S : \text{Sun roof}, P : \text{Power steering}$$

2.14 If $S = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ and $A = \{0, 2, 4, 6, 8\}$, $B = \{1, 3, 5, 7, 9\}$, $C = \{2, 3, 4, 5\}$, and $D = \{1, 6, 7\}$, list the elements of the sets corresponding to the following events:

- (a) $A \cup C$;
 (b) $A \cap B$;
 (c) C' ;
 (d) $(C' \cap D) \cup B$;
 (e) $(S \cap C)'$;
 (f) $A \cap C \cap D'$.

2.15 Consider the sample space $S = \{\text{copper, sodium, nitrogen, potassium, uranium, oxygen, zinc}\}$ and the events

$$\begin{aligned} A &= \{\text{copper, sodium, zinc}\}, \\ B &= \{\text{sodium, nitrogen, potassium}\}, \\ C &= \{\text{oxygen}\}. \end{aligned}$$

List the elements of the sets corresponding to the following events:

- (a) A' ;
 (b) $A \cup C$;
 (c) $(A \cap B') \cup C'$;
 (d) $B' \cap C'$;
 (e) $A \cap B \cap C$;
 (f) $(A' \cup B') \cap (A' \cap C)$.

2.16 If $S = \{x \mid 0 < x < 12\}$, $M = \{x \mid 1 < x < 9\}$, and $N = \{x \mid 0 < x < 5\}$, find

- (a) $M \cup N$;
 (b) $M \cap N$;
 (c) $M' \cap N'$.

2.17 Let A , B , and C be events relative to the sample space S . Using Venn diagrams, shade the areas representing the following events:

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

2.18 Which of the following pairs of events are mutually exclusive?

- (a) A golfer scoring the lowest 18-hole round in a 72-hole tournament and losing the tournament.
- (b) A poker player getting a flush (all cards in the same suit) and 3 of a kind on the same 5-card hand.
- (c) A mother giving birth to a baby girl and a set of twin daughters on the same day.
- (d) A chess player losing the last game and winning the match.

2.19 Suppose that a family is leaving on a summer vacation in their camper and that M is the event that they will experience mechanical problems, T is the event that they will receive a ticket for committing a traffic violation, and V is the event that they will arrive at a campsite with no vacancies. Referring to the Venn diagram of Figure 2.5, state in words the events represented by the following regions:

- (a) region 5;

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

2.20 Referring to Exercise 2.19 and the Venn diagram of Figure 2.5, list the numbers of the regions that represent the following events:

- (a) The family will experience no mechanical problems and will not receive a ticket for a traffic violation but will arrive at a campsite with no vacancies.
- (b) The family will experience both mechanical problems and trouble in locating a campsite with a vacancy but will not receive a ticket for a traffic violation.
- (c) The family will either have mechanical trouble or arrive at a campsite with no vacancies but will not receive a ticket for a traffic violation.
- (d) The family will not arrive at a campsite with no vacancies.

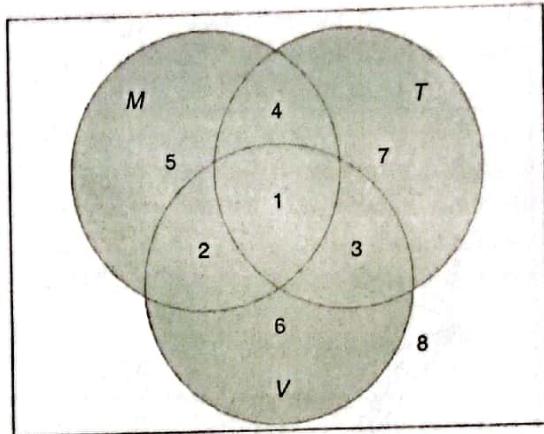


Figure 2.5: Venn diagram for Exercises 2.19 and 2.20.

2.3 Counting Sample Points

One of the problems that the statistician must consider and attempt to evaluate is the element of chance associated with the occurrence of certain events when an experiment is performed. These problems belong in the field of probability, a subject to be introduced in Section 2.4. In many cases, we shall be able to solve a probability problem by counting the number of points in the sample space without actually listing each element. The fundamental principle of counting, often referred to as the multiplication rule, is stated in Rule 2.1.

2.3 Counting Sample Points

Rule 2.1: If an operation can be performed in n_1 ways, and if for each of these ways a second operation can be performed in n_2 ways, then the two operations can be performed together in $n_1 n_2$ ways.

Example 2.13: How many sample points are there in the sample space when a pair of dice is thrown once?

Solution: The first die can land face-up in any one of $n_1 = 6$ ways. For each of these 6 ways, the second die can also land face-up in $n_2 = 6$ ways. Therefore, the pair of dice can land in $n_1 n_2 = (6)(6) = 36$ possible ways.

Example 2.14: A developer of a new subdivision offers prospective home buyers a choice of Tudor, rustic, colonial, and traditional exterior styling in ranch, two-story, and split-level floor plans. In how many different ways can a buyer order one of these homes?

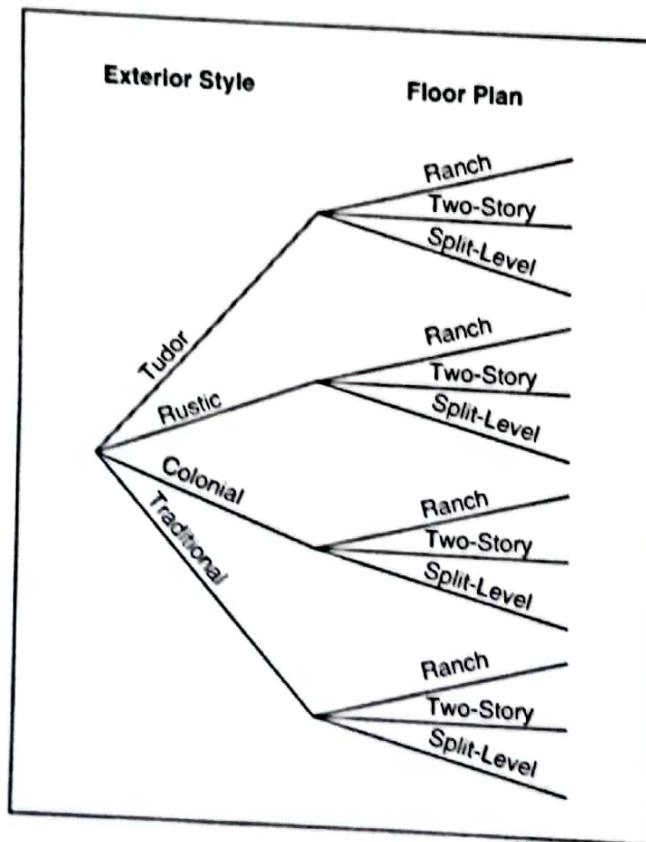


Figure 2.6: Tree diagram for Example 2.14.

Solution: Since $n_1 = 4$ and $n_2 = 3$, a buyer must choose from

$$n_1 n_2 = (4)(3) = 12 \text{ possible homes.}$$

The answers to the two preceding examples can be verified by constructing tree diagrams and counting the various paths along the branches. For instance,

in Example 2.14 there will be $n_1 = 4$ branches corresponding to the different exterior styles, and then there will be $n_2 = 3$ branches extending from each of these 4 branches to represent the different floor plans. This tree diagram yields the $n_1 n_2 = 12$ choices of homes given by the paths along the branches, as illustrated in Figure 2.6.

Example 2.15: If a 22-member club needs to elect a chair and a treasurer, how many different ways can these two be elected?

Solution: For the chair position, there are 22 total possibilities. For each of those 22 possibilities, there are 21 possibilities to elect the treasurer. Using the multiplication rule, we obtain $n_1 \times n_2 = 22 \times 21 = 462$ different ways.

The multiplication rule, Rule 2.1 may be extended to cover any number of operations. Suppose, for instance, that a customer wishes to buy a new cell phone and can choose from $n_1 = 5$ brands, $n_2 = 5$ sets of capability, and $n_3 = 4$ colors. These three classifications result in $n_1 n_2 n_3 = (5)(5)(4) = 100$ different ways for a customer to order one of these phones. The **generalized multiplication rule** covering k operations is stated in the following.

Rule 2.2: If an operation can be performed in n_1 ways, and if for each of these a second operation can be performed in n_2 ways, and for each of the first two a third operation can be performed in n_3 ways, and so forth, then the sequence of k operations can be performed in $n_1 n_2 \cdots n_k$ ways.

Example 2.16: Sam is going to assemble a computer by himself. He has the choice of chips from two brands, a hard drive from four, memory from three, and an accessory bundle from five local stores. How many different ways can Sam order the parts?

Solution: Since $n_1 = 2$, $n_2 = 4$, $n_3 = 3$, and $n_4 = 5$, there are

$$n_1 \times n_2 \times n_3 \times n_4 = 2 \times 4 \times 3 \times 5 = 120$$

different ways to order the parts.

Example 2.17: How many even four-digit numbers can be formed from the digits 0, 1, 2, 5, 6, and 9 if each digit can be used only once?

Solution: Since the number must be even, we have only $n_1 = 3$ choices for the units position. However, for a four-digit number the thousands position cannot be 0. Hence, we consider the units position in two parts, 0 or not 0. If the units position is 0 (i.e., $n_1 = 1$), we have $n_2 = 5$ choices for the thousands position, $n_3 = 4$ for the hundreds position, and $n_4 = 3$ for the tens position. Therefore, in this case we have a total of

$$n_1 n_2 n_3 n_4 = (1)(5)(4)(3) = 60$$

even four-digit numbers. On the other hand, if the units position is not 0 (i.e., $n_1 = 2$), we have $n_2 = 4$ choices for the thousands position, $n_3 = 4$ for the hundreds position, and $n_4 = 3$ for the tens position. In this situation, there are a total of

$$n_1 n_2 n_3 n_4 = (2)(4)(4)(3) = 96$$

2.3 Counting Sample Points

even four-digit numbers.

Since the above two cases are mutually exclusive, the total number of even four-digit numbers can be calculated as $60 + 96 = 156$.

Frequently, we are interested in a sample space that contains as elements all possible orders or arrangements of a group of objects. For example, we may want to know how many different arrangements are possible for sitting 6 people around a table, or we may ask how many different orders are possible for drawing 2 lottery tickets from a total of 20. The different arrangements are called permutations.

Definition 2.7: A permutation is an arrangement of all or part of a set of objects.

Consider the three letters a , b , and c . The possible permutations are abc , acb , bac , bca , cab , and cba . Thus, we see that there are 6 distinct arrangements. Using Rule 2.2, we could arrive at the answer 6 without actually listing the different orders by the following arguments: There are $n_1 = 3$ choices for the first position. No matter which letter is chosen, there are always $n_2 = 2$ choices for the second position. No matter which two letters are chosen for the first two positions, there is only $n_3 = 1$ choice for the last position, giving a total of

$$n_1 n_2 n_3 = (3)(2)(1) = 6 \text{ permutations}$$

by Rule 2.2. In general, n distinct objects can be arranged in

$$n(n - 1)(n - 2) \cdots (3)(2)(1) \text{ ways.}$$

There is a notation for such a number.

Definition 2.8: For any non-negative integer n , $n!$, called " n factorial," is defined as

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

with special case $0! = 1$.

Using the argument above, we arrive at the following theorem.

Theorem 2.1: The number of permutations of n objects is $n!$.

The number of permutations of the four letters a , b , c , and d will be $4! = 24$. Now consider the number of permutations that are possible by taking two letters at a time from four. These would be ab , ac , ad , ba , bc , bd , ca , cb , cd , da , db , and dc . Using Rule 2.1 again, we have two positions to fill, with $n_1 = 4$ choices for the first and then $n_2 = 3$ choices for the second, for a total of

$$n_1 n_2 = (4)(3) = 12$$

permutations. In general, n distinct objects taken r at a time can be arranged in

$$n(n - 1)(n - 2) \cdots (n - r + 1)$$

ways. We represent this product by the symbol

$${}_nP_r = \frac{n!}{(n - r)!}.$$

As a result, we have the theorem that follows.

Theorem 2.2: The number of permutations of n distinct objects taken r at a time is

$${}_n P_r = \frac{n!}{(n - r)!}$$

Example 2.18: In one year, three awards (research, teaching, and service) will be given to a class of 25 graduate students in a statistics department. If each student can receive at most one award, how many possible selections are there?

Solution: Since the awards are distinguishable, it is a permutation problem. The total number of sample points is

$${}_{25} P_3 = \frac{25!}{(25 - 3)!} = \frac{25!}{22!} = (25)(24)(23) = 13,800.$$

Example 2.19: A president and a treasurer are to be chosen from a student club consisting of 50 people. How many different choices of officers are possible if

- (a) there are no restrictions;
- (b) A will serve only if he is president;
- (c) B and C will serve together or not at all;
- (d) D and E will not serve together?

Solution: (a) The total number of choices of officers, without any restrictions, is

$${}_{50} P_2 = \frac{50!}{48!} = (50)(49) = 2450.$$

- (b) Since A will serve only if he is president, we have two situations here: (i) A is selected as the president, which yields 49 possible outcomes for the treasurer's position, or (ii) officers are selected from the remaining 49 people without A , which has the number of choices ${}_{49} P_2 = (49)(48) = 2352$. Therefore, the total number of choices is $49 + 2352 = 2401$.
- (c) The number of selections when B and C serve together is 2. The number of selections when both B and C are not chosen is ${}_{48} P_2 = 2256$. Therefore, the total number of choices in this situation is $2 + 2256 = 2258$.
- (d) The number of selections when D serves as an officer but not E is $(2)(48) = 96$, where 2 is the number of positions D can take and 48 is the number of selections of the other officer from the remaining people in the club except E . The number of selections when E serves as an officer but not D is also $(2)(48) = 96$. The number of selections when both D and E are not chosen is ${}_{48} P_2 = 2256$. Therefore, the total number of choices is $(2)(96) + 2256 = 2448$. This problem also has another short solution: Since D and E can only serve together in 2 ways, the answer is $2450 - 2 = 2448$.

2.3 Counting Sample Points

Permutations that occur by arranging objects in a circle are called **circular permutations**. Two circular permutations are not considered different unless corresponding objects in the two arrangements are preceded or followed by a different object as we proceed in a clockwise direction. For example, if 4 people are playing bridge, we do not have a new permutation if they all move one position in a clockwise direction. By considering one person in a fixed position and arranging the other three in $3!$ ways, we find that there are 6 distinct arrangements for the bridge game.

Theorem 2.3: The number of permutations of n objects arranged in a circle is $(n - 1)!$.

So far we have considered permutations of distinct objects. That is, all the objects were completely different or distinguishable. Obviously, if the letters b and c are both equal to x , then the 6 permutations of the letters a , b , and c become axx , axx , rax , rax , xxa , and xxa , of which only 3 are distinct. Therefore, with 3 letters, 2 being the same, we have $3!/2! = 3$ distinct permutations. With 4 different letters a , b , c , and d , we have 24 distinct permutations. If we let $a = b = x$ and $c = d = y$, we can list only the following distinct permutations: $xxyy$, $xyxy$, yxx , $yyxx$, $xyyx$, and $yxyx$. Thus, we have $4!/(2! 2!) = 6$ distinct permutations.

Theorem 2.4: The number of distinct permutations of n things of which n_1 are of one kind, n_2 of a second kind, ..., n_k of a k th kind is

$$\frac{n!}{n_1! n_2! \cdots n_k!}.$$

Example 2.20: In a college football training session, the defensive coordinator needs to have 10 players standing in a row. Among these 10 players, there are 1 freshman, 2 sophomores, 4 juniors, and 3 seniors. How many different ways can they be arranged in a row if only their class level will be distinguished?

Solution: Directly using Theorem 2.4, we find that the total number of arrangements is

$$\frac{10!}{1! 2! 4! 3!} = 12,600.$$

Often we are concerned with the number of ways of partitioning a set of n objects into r subsets called **cells**. A partition has been achieved if the intersection of every possible pair of the r subsets is the empty set ϕ and if the union of all subsets gives the original set. The order of the elements within a cell is of no importance. Consider the set $\{a, e, i, o, u\}$. The possible partitions into two cells in which the first cell contains 4 elements and the second cell 1 element are

$$\{(a, e, i, o), (u)\}, \{(a, i, o, u), (e)\}, \{(e, i, o, u), (a)\}, \{(a, e, o, u), (i)\}, \{(a, e, i, u), (o)\}.$$

We see that there are 5 ways to partition a set of 4 elements into two subsets, or cells, containing 4 elements in the first cell and 1 element in the second.

The number of partitions for this illustration is denoted by the symbol

$$\binom{5}{4,1} = \frac{5!}{4! 1!} = 5,$$

where the top number represents the total number of elements and the bottom numbers represent the number of elements going into each cell. We state this more generally in Theorem 2.5.

Theorem 2.5:

The number of ways of partitioning a set of n objects into r cells with n_1 elements in the first cell, n_2 elements in the second, and so forth, is

$$\binom{n}{n_1, n_2, \dots, n_r} = \frac{n!}{n_1! n_2! \cdots n_r!},$$

where $n_1 + n_2 + \cdots + n_r = n$.

Example 2.21: In how many ways can 7 graduate students be assigned to 1 triple and 2 double hotel rooms during a conference?

Solution: The total number of possible partitions would be

$$\binom{7}{3, 2, 2} = \frac{7!}{3! 2! 2!} = 210.$$

In many problems, we are interested in the number of ways of selecting r objects from n without regard to order. These selections are called **combinations**. A combination is actually a partition with two cells, the one cell containing the r objects selected and the other cell containing the $(n - r)$ objects that are left. The number of such combinations, denoted by

$$\binom{n}{r, n-r}, \text{ is usually shortened to } \binom{n}{r},$$

since the number of elements in the second cell must be $n - r$.

Theorem 2.6:

The number of combinations of n distinct objects taken r at a time is

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

Example 2.22: A young boy asks his mother to get 5 Game-Boy™ cartridges from his collection of 10 arcade and 5 sports games. How many ways are there that his mother can get 3 arcade and 2 sports games?

Solution: The number of ways of selecting 3 cartridges from 10 is

$$\binom{10}{3} = \frac{10!}{3! (10-3)!} = 120.$$

The number of ways of selecting 2 cartridges from 5 is

$$\binom{5}{2} = \frac{5!}{2! 3!} = 10.$$

Exercises

Using the multiplication rule (Rule 2.1) with $n_1 = 120$ and $n_2 = 10$, we have
 $(120)(10) = 1200$ ways.

Example 2.23: How many different letter arrangements can be made from the letters in the word **STATISTICS?**

Solution: Using the same argument as in the discussion for Theorem 2.6, in this example we can actually apply Theorem 2.5 to obtain

$$\binom{10}{3, 3, 2, 1, 1} = \frac{10!}{3! 3! 2! 1! 1!} = 50,400.$$

Here we have 10 total letters, with 2 letters (*S*, *T*) appearing 3 times each, letter *I* appearing twice, and letters *A* and *C* appearing once each. On the other hand, this result can be directly obtained by using Theorem 2.4. ■

Exercises

2.21 Registrants at a large convention are offered 6 sightseeing tours on each of 3 days. In how many ways can a person arrange to go on a sightseeing tour planned by this convention?

2.22 In a medical study, patients are classified in 8 ways according to whether they have blood type AB^+ , AB^- , A^+ , A^- , B^+ , B^- , O^+ , or O^- , and also according to whether their blood pressure is low, normal, or high. Find the number of ways in which a patient can be classified.

2.23 If an experiment consists of throwing a die and then drawing a letter at random from the English alphabet, how many points are there in the sample space?

2.24 Students at a private liberal arts college are classified as being freshmen, sophomores, juniors, or seniors, and also according to whether they are male or female. Find the total number of possible classifications for the students of that college.

2.25 A certain brand of shoes comes in 5 different styles, with each style available in 4 distinct colors. If the store wishes to display pairs of these shoes showing all of its various styles and colors, how many different pairs will the store have on display?

2.26 A California study concluded that following 7 simple health rules can extend a man's life by 11 years on the average and a woman's life by 7 years. These 7 rules are as follows: no smoking, get regular exercise, use alcohol only in moderation, get 7 to 8 hours of sleep, maintain proper weight, eat breakfast, and do

not eat between meals. In how many ways can a person adopt 5 of these rules to follow

- (a) if the person presently violates all 7 rules?
- (b) if the person never drinks and always eats breakfast?

2.27 A developer of a new subdivision offers a prospective home buyer a choice of 4 designs, 3 different heating systems, a garage or carport, and a patio or screened porch. How many different plans are available to this buyer?

2.28 A drug for the relief of asthma can be purchased from 5 different manufacturers in liquid, tablet, or capsule form, all of which come in regular and extra strength. How many different ways can a doctor prescribe the drug for a patient suffering from asthma?

2.29 In a fuel economy study, each of 3 race cars is tested using 5 different brands of gasoline at 7 test sites located in different regions of the country. If 2 drivers are used in the study, and test runs are made once under each distinct set of conditions, how many test runs are needed?

2.30 In how many different ways can a true-false test consisting of 9 questions be answered?

2.31 A witness to a hit-and-run accident told the police that the license number contained the letters RLH followed by 3 digits, the first of which was a 5. If the witness cannot recall the last 2 digits, but is certain that all 3 digits are different, find the maximum number of automobile registrations that the police may have to check.

- 2.32** (a) In how many ways can 6 people be lined up to get on a bus?
 (b) If 3 specific persons, among 6, insist on following each other, how many ways are possible?
 (c) If 2 specific persons, among 6, refuse to follow each other, how many ways are possible?
- 2.33** If a multiple-choice test consists of 5 questions, each with 4 possible answers of which only 1 is correct,
 (a) in how many different ways can a student check off one answer to each question?
 (b) in how many ways can a student check off one answer to each question and get all the answers wrong?
- 2.34** (a) How many distinct permutations can be made from the letters of the word *COLUMNS*?
 (b) How many of these permutations start with the letter *M*?
- 2.35** A contractor wishes to build 9 houses, each different in design. In how many ways can he place these houses on a street if 6 lots are on one side of the street and 3 lots are on the opposite side?
- 2.36** (a) How many three-digit numbers can be formed from the digits 0, 1, 2, 3, 4, 5, and 6 if each digit can be used only once?
 (b) How many of these are odd numbers?
 (c) How many are greater than 330?
- 2.37** In how many ways can 4 boys and 5 girls sit in a row if the boys and girls must alternate?
- 2.38** Four married couples have bought 8 seats in the same row for a concert. In how many different ways can they be seated
 (a) with no restrictions?
 (b) if each couple is to sit together?
- (c) if all the men sit together to the right of all the women?
- 2.39** In a regional spelling bee, the 8 finalists consist of 3 boys and 5 girls. Find the number of sample points in the sample space S for the number of possible orders at the conclusion of the contest for
 (a) all 8 finalists;
 (b) the first 3 positions.
- 2.40** In how many ways can 5 starting positions on a basketball team be filled with 8 men who can play any of the positions?
- 2.41** Find the number of ways that 6 teachers can be assigned to 4 sections of an introductory psychology course if no teacher is assigned to more than one section.
- 2.42** Three lottery tickets for first, second, and third prizes are drawn from a group of 40 tickets. Find the number of sample points in S for awarding the 3 prizes if each contestant holds only 1 ticket.
- 2.43** In how many ways can 5 different trees be planted in a circle?
- 2.44** In how many ways can a caravan of 8 covered wagons from Arizona be arranged in a circle?
- 2.45** How many distinct permutations can be made from the letters of the word *INFINITY*?
- 2.46** In how many ways can 3 oaks, 4 pines, and 2 maples be arranged along a property line if one does not distinguish among trees of the same kind?
- 2.47** How many ways are there to select 3 candidates from 8 equally qualified recent graduates for openings in an accounting firm?
- 2.48** How many ways are there that no two students will have the same birth date in a class of size 60?

2.4 Probability of an Event

Perhaps it was humankind's unquenchable thirst for gambling that led to the early development of probability theory. In an effort to increase their winnings, gamblers called upon mathematicians to provide optimum strategies for various games of chance. Some of the mathematicians providing these strategies were Pascal, Leibniz, Fermat, and James Bernoulli. As a result of this development of probability theory, statistical inference, with all its predictions and generalizations, has branched out far beyond games of chance to encompass many other fields associated with chance occurrences, such as politics, business, weather forecasting,

2.4 Probability of an Event

and scientific research. For these predictions and generalizations to be reasonably accurate, an understanding of basic probability theory is essential.

What do we mean when we make the statement "John will probably win the tennis match," or "I have a fifty-fifty chance of getting an even number when a die is tossed," or "The university is not likely to win the football game tonight," or "Most of our graduating class will likely be married within 3 years"? In each case, we are expressing an outcome of which we are not certain, but owing to past information or from an understanding of the structure of the experiment, we have some degree of confidence in the validity of the statement.

Throughout the remainder of this chapter, we consider only those experiments for which the sample space contains a finite number of elements. The likelihood of the occurrence of an event resulting from such a statistical experiment is evaluated by means of a set of real numbers, called **weights** or **probabilities**, ranging from 0 to 1. To every point in the sample space we assign a probability such that the sum of all probabilities is 1. If we have reason to believe that a certain sample point is quite likely to occur when the experiment is conducted, the probability assigned should be close to 1. On the other hand, a probability closer to 0 is assigned to a sample point that is not likely to occur. In many experiments, such as tossing a coin or a die, all the sample points have the same chance of occurring and are assigned equal probabilities. For points outside the sample space, that is, for simple events that cannot possibly occur, we assign a probability of 0.

To find the probability of an event A , we sum all the probabilities assigned to the sample points in A . This sum is called the **probability of A** and is denoted by $P(A)$.

Definition 2.9:

The probability of an event A is the sum of the weights of all sample points in A . Therefore,

$$0 \leq P(A) \leq 1, \quad P(\emptyset) = 0, \quad \text{and} \quad P(S) = 1.$$

Furthermore, if A_1, A_2, A_3, \dots is a sequence of mutually exclusive events, then

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

Example 2.24: A coin is tossed twice. What is the probability that at least 1 head occurs?

Solution: The sample space for this experiment is

$$S = \{HH, HT, TH, TT\}.$$

If the coin is balanced, each of these outcomes is equally likely to occur. Therefore, we assign a probability of ω to each sample point. Then $4\omega = 1$, or $\omega = 1/4$. If A represents the event of at least 1 head occurring, then

$$A = \{HH, HT, TH\} \text{ and } P(A) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}.$$

Example 2.25: A die is loaded in such a way that an even number is twice as likely to occur as an odd number. If E is the event that a number less than 4 occurs on a single toss of the die, find $P(E)$.

Solution: The sample space is $S = \{1, 2, 3, 4, 5, 6\}$. We assign a probability of w to each odd number and a probability of $2w$ to each even number. Since the sum of the probabilities must be 1, we have $9w = 1$ or $w = 1/9$. Hence, probabilities of $1/9$ and $2/9$ are assigned to each odd and even number, respectively. Therefore,

$$E = \{1, 2, 3\} \text{ and } P(E) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}.$$

Example 2.26: In Example 2.25, let A be the event that an even number turns up and let B be the event that a number divisible by 3 occurs. Find $P(A \cup B)$ and $P(A \cap B)$.

Solution: For the events $A = \{2, 4, 6\}$ and $B = \{3, 6\}$, we have

$$A \cup B = \{2, 3, 4, 6\} \text{ and } A \cap B = \{6\}.$$

By assigning a probability of $1/9$ to each odd number and $2/9$ to each even number, we have

$$P(A \cup B) = \frac{2}{9} + \frac{1}{9} + \frac{2}{9} + \frac{2}{9} = \frac{7}{9} \quad \text{and} \quad P(A \cap B) = \frac{2}{9}.$$

If the sample space for an experiment contains N elements, all of which are equally likely to occur, we assign a probability equal to $1/N$ to each of the N points. The probability of any event A containing n of these N sample points is then the ratio of the number of elements in A to the number of elements in S .

Rule 2.3: If an experiment can result in any one of N different equally likely outcomes, and if exactly n of these outcomes correspond to event A , then the probability of event A is

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

Example 2.27: A statistics class for engineers consists of 25 industrial, 10 mechanical, 10 electrical, and 8 civil engineering students. If a person is randomly selected by the instructor to answer a question, find the probability that the student chosen is (a) an industrial engineering major and (b) a civil engineering or an electrical engineering major.

Solution: Denote by I , M , E , and C the students majoring in industrial, mechanical, electrical, and civil engineering, respectively. The total number of students in the class is 53, all of whom are equally likely to be selected.

- (a) Since 25 of the 53 students are majoring in industrial engineering, the probability of event I , selecting an industrial engineering major at random, is

$$P(I) = \frac{25}{53}.$$

- (b) Since 18 of the 53 students are civil or electrical engineering majors, it follows that

$$P(C \cup E) = \frac{18}{53}.$$

Example 2.28: In a poker hand consisting of 5 cards, find the probability of holding 2 aces and 3 jacks.

Solution: The number of ways of being dealt 2 aces from 4 cards is

$$\binom{4}{2} = \frac{4!}{2! 2!} = 6,$$

and the number of ways of being dealt 3 jacks from 4 cards is

$$\binom{4}{3} = \frac{4!}{3! 1!} = 4.$$

By the multiplication rule (Rule 2.1), there are $n = (6)(4) = 24$ hands with 2 aces and 3 jacks. The total number of 5-card poker hands, all of which are equally likely, is

$$N = \binom{52}{5} = \frac{52!}{5! 47!} = 2,598,960.$$

Therefore, the probability of getting 2 aces and 3 jacks in a 5-card poker hand is

$$P(C) = \frac{24}{2,598,960} = 0.9 \times 10^{-5}. \quad \blacksquare$$

If the outcomes of an experiment are not equally likely to occur, the probabilities must be assigned on the basis of prior knowledge or experimental evidence. For example, if a coin is not balanced, we could estimate the probabilities of heads and tails by tossing the coin a large number of times and recording the outcomes. According to the **relative frequency** definition of probability, the true probabilities would be the fractions of heads and tails that occur in the long run. Another intuitive way of understanding probability is the **indifference** approach. For instance, if you have a die that you believe is balanced, then using this indifference approach, you determine that the probability that each of the six sides will show up after a throw is $1/6$.

To find a numerical value that represents adequately the probability of winning at tennis, we must depend on our past performance at the game as well as that of the opponent and, to some extent, our belief in our ability to win. Similarly, to find the probability that a horse will win a race, we must arrive at a probability based on the previous records of all the horses entered in the race as well as the records of the jockeys riding the horses. Intuition would undoubtedly also play a part in determining the size of the bet that we might be willing to wager. The use of intuition, personal beliefs, and other indirect information in arriving at probabilities is referred to as the **subjective** definition of probability.

In most of the applications of probability in this book, the relative frequency interpretation of probability is the operative one. Its foundation is the statistical experiment rather than subjectivity, and it is best viewed as the **limiting relative frequency**. As a result, many applications of probability in science and engineering must be based on experiments that can be repeated. Less objective notions of probability are encountered when we assign probabilities based on prior information and opinions, as in "There is a good chance that the Giants will lose the Super

Bowl." When opinions and prior information differ from individual to individual, subjective probability becomes the relevant resource. In Bayesian statistics (see Chapter 18), a more subjective interpretation of probability will be used, based on an elicitation of prior probability information.

2.5 Additive Rules

Often it is easiest to calculate the probability of some event from known probabilities of other events. This may well be true if the event in question can be represented as the union of two other events or as the complement of some event. Several important laws that frequently simplify the computation of probabilities follow. The first, called the **additive rule**, applies to unions of events.

Theorem 2.7: If A and B are two events, then

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

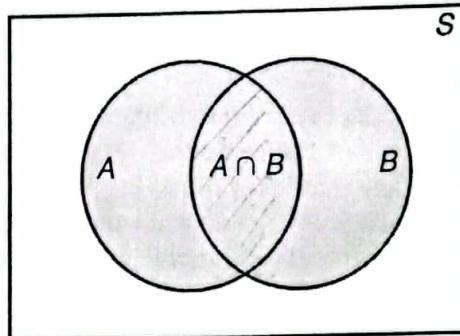
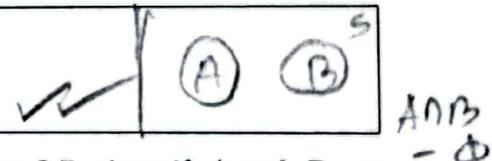


Figure 2.7: Additive rule of probability.

Proof: Consider the Venn diagram in Figure 2.7. The $P(A \cup B)$ is the sum of the probabilities of the sample points in $A \cup B$. Now $P(A) + P(B)$ is the sum of all the probabilities in A plus the sum of all the probabilities in B . Therefore, we have added the probabilities in $(A \cap B)$ twice. Since these probabilities add up to $P(A \cap B)$, we must subtract this probability once to obtain the sum of the probabilities in $A \cup B$. ■

Corollary 2.1: If A and B are mutually exclusive, then

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



Corollary 2.1 is an immediate result of Theorem 2.7, since if A and B are mutually exclusive, $A \cap B = \emptyset$ and then $P(A \cap B) = P(\emptyset) = 0$. In general, we can write Corollary 2.2.

Corollary 2.2: If A_1, A_2, \dots, A_n are mutually exclusive, then

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n).$$

A collection of events $\{A_1, A_2, \dots, A_n\}$ of a sample space S is called a **partition** of S if A_1, A_2, \dots, A_n are mutually exclusive and $A_1 \cup A_2 \cup \dots \cup A_n = S$. Thus, we have

Corollary 2.3: If A_1, A_2, \dots, A_n is a partition of sample space S , then

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n) = P(S) = 1.$$

As one might expect, Theorem 2.7 extends in an analogous fashion.

Theorem 2.8: For three events A, B , and C ,

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

Example 2.29: John is going to graduate from an industrial engineering department in a university by the end of the semester. After being interviewed at two companies he likes, he assesses that his probability of getting an offer from company A is 0.8, and his probability of getting an offer from company B is 0.6. If he believes that the probability that he will get offers from both companies is 0.5, what is the probability that he will get at least one offer from these two companies?

Solution: Using the additive rule, we have

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.8 + 0.6 - 0.5 = 0.9.$$

Example 2.30: What is the probability of getting a total of 7 or 11 when a pair of fair dice is tossed?

Solution: Let A be the event that 7 occurs and B the event that 11 comes up. Now, a total of 7 occurs for 6 of the 36 sample points, and a total of 11 occurs for only 2 of the sample points. Since all sample points are equally likely, we have $P(A) = 1/6$ and $P(B) = 1/18$. The events A and B are mutually exclusive, since a total of 7 and 11 cannot both occur on the same toss. Therefore,

$$P(A \cup B) = P(A) + P(B) = \frac{1}{6} + \frac{1}{18} = \frac{2}{9}.$$

This result could also have been obtained by counting the total number of points for the event $A \cup B$, namely 8, and writing

$$P(A \cup B) = \frac{n}{N} = \frac{8}{36} = \frac{2}{9}.$$

Theorem 2.7 and its three corollaries should help the reader gain more insight into probability and its interpretation. Corollaries 2.1 and 2.2 suggest the very intuitive result dealing with the probability of occurrence of at least one of a number of events, no two of which can occur simultaneously. The probability that at least one occurs is the sum of the probabilities of occurrence of the individual events. The third corollary simply states that the highest value of a probability (unity) is assigned to the entire sample space S .

Example 2.31: If the probabilities are, respectively, 0.09, 0.15, 0.21, and 0.23 that a person purchasing a new automobile will choose the color green, white, red, or blue, what is the probability that a given buyer will purchase a new automobile that comes in one of those colors?

Solution: Let G , W , R , and B be the events that a buyer selects, respectively, a green, white, red, or blue automobile. Since these four events are mutually exclusive, the probability is

$$\begin{aligned} P(G \cup W \cup R \cup B) &= P(G) + P(W) + P(R) + P(B) \\ &= 0.09 + 0.15 + 0.21 + 0.23 = 0.68. \end{aligned}$$

Often it is more difficult to calculate the probability that an event occurs than it is to calculate the probability that the event does not occur. Should this be the case for some event A , we simply find $P(A')$ first and then, using Theorem 2.7, find $P(A)$ by subtraction.

Theorem 2.9: If A and A' are complementary events, then

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

Proof: Since $A \cup A' = S$ and the sets A and A' are disjoint,

$$1 = P(S) = P(A \cup A') = P(A) + P(A').$$

Example 2.32: If the probabilities that an automobile mechanic will service 3, 4, 5, 6, 7, or 8 or more cars on any given workday are, respectively, 0.12, 0.19, 0.28, 0.24, 0.10, and 0.07, what is the probability that he will service at least 5 cars on his next day at work?

Solution: Let E be the event that at least 5 cars are serviced. Now, $P(E) = 1 - P(E')$, where E' is the event that fewer than 5 cars are serviced. Since

$$P(E') = 0.12 + 0.19 = 0.31,$$

it follows from Theorem 2.9 that

$$P(E) = 1 - 0.31 = 0.69.$$

Example 2.33: Suppose the manufacturer's specifications for the length of a certain type of computer cable are 2000 ± 10 millimeters. In this industry, it is known that small cable is just as likely to be defective (not meeting specifications) as large cable. That is,

If it will locate in D, the probability that it will be located in D is 0.8. If it will locate in E, the probability that it will be located in E is 0.2.

the probability of randomly producing a cable with length exceeding 2010 millimeters is equal to the probability of producing a cable with length smaller than 1990 millimeters. The probability that the production procedure meets specifications is known to be 0.99.

- (a) What is the probability that a cable selected randomly is too large?
- (b) What is the probability that a randomly selected cable is larger than 1990 millimeters?

Solution: Let M be the event that a cable meets specifications. Let S and L be the events that the cable is too small and too large, respectively. Then

- (a) $P(M) = 0.99$ and $P(S) = P(L) = (1 - 0.99)/2 = 0.005$.
- (b) Denoting by X the length of a randomly selected cable, we have

$$P(1990 \leq X \leq 2010) = P(M) = 0.99.$$

Since $P(X \geq 2010) = P(L) = 0.005$,

$$P(X \geq 1990) = P(M) + P(L) = 0.995.$$

This also can be solved by using Theorem 2.9:

$$P(X \geq 1990) + P(X < 1990) = 1.$$

Thus, $P(X \geq 1990) = 1 - P(S) = 1 - 0.005 = 0.995$. ■

Exercises

2.49 Find the errors in each of the following statements:

- (a) The probabilities that an automobile salesperson will sell 0, 1, 2, or 3 cars on any given day in February are, respectively, 0.19, 0.38, 0.29, and 0.15.
- (b) The probability that it will rain tomorrow is 0.40, and the probability that it will not rain tomorrow is 0.52.
- (c) The probabilities that a printer will make 0, 1, 2, 3, or 4 or more mistakes in setting a document are, respectively, 0.19, 0.34, -0.25, 0.43, and 0.29.
- (d) On a single draw from a deck of playing cards, the probability of selecting a heart is $1/4$, the probability of selecting a black card is $1/2$, and the probability of selecting both a heart and a black card is $1/8$.

2.50 Assuming that all elements of S in Exercise 2.8 on page 42 are equally likely to occur, find

- (a) the probability of event A ;
- (b) the probability of event C ;
- (c) the probability of event $A \cap C$.

2.51 A box contains 500 envelopes, of which 75 contain \$100 in cash, 150 contain \$25, and 275 contain \$10. An envelope may be purchased for \$25. What is the sample space for the different amounts of money? Assign probabilities to the sample points and then find the probability that the first envelope purchased contains less than \$100.

2.52 Suppose that in a senior college class of 500 students it is found that 210 smoke, 258 drink alcoholic beverages, 216 eat between meals, 122 smoke and drink alcoholic beverages, 83 eat between meals and drink alcoholic beverages, 97 smoke and eat between meals, and 52 engage in all three of these bad health practices. If a member of this senior class is selected at random, find the probability that the student

- (a) smokes but does not drink alcoholic beverages;
- (b) eats between meals and drinks alcoholic beverages but does not smoke;
- (c) neither smokes nor eats between meals.

2.53 The probability that an American industry will locate in Shanghai, China, is 0.7, the probability that

it will locate in Beijing, China, is 0.4, and the probability that it will locate in either Shanghai or Beijing or both is 0.8. What is the probability that the industry will locate?

- (a) in both cities?
- (b) in neither city?

2.54 From past experience, a stockbroker believes that under present economic conditions a customer will invest in tax-free bonds with a probability of 0.6, will invest in mutual funds with a probability of 0.3, and will invest in both tax-free bonds and mutual funds with a probability of 0.15. At this time, find the probability that a customer will invest

- (a) in either tax-free bonds or mutual funds;
- (b) in neither tax-free bonds nor mutual funds.

2.55 If each coded item in a catalog begins with 3 distinct letters followed by 4 distinct nonzero digits, find the probability of randomly selecting one of these coded items with the first letter a vowel and the last digit even

2.56 An automobile manufacturer is concerned about a possible recall of its best-selling four-door sedan. If there were a recall, there is a probability of 0.25 of a defect in the brake system, 0.18 of a defect in the transmission, 0.17 of a defect in the fuel system, and 0.40 of a defect in some other area.

- (a) What is the probability that the defect is the brakes or the fueling system if the probability of defects in both systems simultaneously is 0.15?
- (b) What is the probability that there are no defects in either the brakes or the fueling system?

2.57 If a letter is chosen at random from the English alphabet, find the probability that the letter

- (a) is a vowel exclusive of y;
- (b) is listed somewhere ahead of the letter j;
- (c) is listed somewhere after the letter g.

2.58 A pair of fair dice is tossed. Find the probability of getting

- (a) a total of 8;
- (b) at most a total of 5.

2.59 In a poker hand consisting of 5 cards, find the probability of holding

- (a) 3 aces;
- (b) 4 hearts and 1 club.

2.60 If 3 books are picked at random from a shelf containing 5 novels, 3 books of poems, and a dictionary, what is the probability that

- (a) the dictionary is selected?
- (b) 2 novels and 1 book of poems are selected?

2.61 In a high school graduating class of 100 students, 54 studied mathematics, 69 studied history, and 35 studied both mathematics and history. If one of these students is selected at random, find the probability that

- (a) the student took mathematics or history;
- (b) the student did not take either of these subjects;
- (c) the student took history but not mathematics.

2.62 Dom's Pizza Company uses taste testing and statistical analysis of the data prior to marketing any new product. Consider a study involving three types of crusts (thin, thin with garlic and oregano, and thin with bits of cheese). Dom's is also studying three sauces (standard, a new sauce with more garlic, and a new sauce with fresh basil).

- (a) How many combinations of crust and sauce are involved?
- (b) What is the probability that a judge will get a plain thin crust with a standard sauce for his first taste test?

2.63 According to *Consumer Digest* (July/August 1996), the probable location of personal computers (PC) in the home is as follows:

Adult bedroom:	0.03
Child bedroom:	0.15
Other bedroom:	0.14
Office or den:	0.40
Other rooms:	0.28

- (a) What is the probability that a PC is in a bedroom?
- (b) What is the probability that it is not in a bedroom?
- (c) Suppose a household is selected at random from households with a PC; in what room would you expect to find a PC?

2.64 Interest centers around the life of an electronic component. Suppose it is known that the probability that the component survives for more than 6000 hours is 0.42. Suppose also that the probability that the component survives *no longer than* 4000 hours is 0.04.

- (a) What is the probability that the life of the component is less than or equal to 6000 hours?
- (b) What is the probability that the life is greater than 4000 hours?

Exercises

2.65 Consider the situation of Exercise 2.64. Let A be the event that the component fails a particular test and B be the event that the component displays strain but does not actually fail. Event A occurs with probability 0.20, and event B occurs with probability 0.35.

- (a) What is the probability that the component does not fail the test?
- (b) What is the probability that the component works perfectly well (i.e., neither displays strain nor fails the test)?
- (c) What is the probability that the component either fails or shows strain in the test?

2.66 Factory workers are constantly encouraged to practice zero tolerance when it comes to accidents in factories. Accidents can occur because the working environment or conditions themselves are unsafe. On the other hand, accidents can occur due to carelessness or so-called human error. In addition, the worker's shift, 7:00 A.M.-3:00 P.M. (day shift), 3:00 P.M.-11:00 P.M. (evening shift), or 11:00 P.M.-7:00 A.M. (graveyard shift), may be a factor. During the last year, 300 accidents have occurred. The percentages of the accidents for the condition combinations are as follows.

Shift	Unsafe Conditions	Human Error
Day	5%	32%
Evening	6%	25%
Graveyard	2%	30%

If an accident report is selected randomly from the 300 reports,

- (a) what is the probability that the accident occurred on the graveyard shift?
- (b) what is the probability that the accident occurred due to human error?
- (c) what is the probability that the accident occurred due to unsafe conditions?
- (d) what is the probability that the accident occurred on either the evening or the graveyard shift?

2.67 Consider the situation of Example 2.32 on page 58.

- (a) What is the probability that no more than 4 cars will be serviced by the mechanic?
- (b) What is the probability that he will service fewer than 8 cars?
- (c) What is the probability that he will service either 3 or 4 cars?

2.68 Interest centers around the nature of an oven purchased at a particular department store. It can be either a gas or an electric oven. Consider the decisions made by six distinct customers.

- (a) Suppose that the probability is 0.40 that at most

two of these individuals purchase an electric oven. What is the probability that at least three purchase the electric oven?

- (b) Suppose it is known that the probability that all six purchase the electric oven is 0.007 while 0.104 is the probability that all six purchase the gas oven. What is the probability that at least one of each type is purchased?

2.69 It is common in many industrial areas to use a filling machine to fill boxes full of product. This occurs in the food industry as well as other areas in which the product is used in the home, for example, detergent. These machines are not perfect, and indeed they may A , fill to specification, B , underfill, and C , overfill. Generally, the practice of underfilling is that which one hopes to avoid. Let $P(B) = 0.001$ while $P(A) = 0.990$.

- (a) Give $P(C)$.
- (b) What is the probability that the machine does not underfill?
- (c) What is the probability that the machine either overfills or underfills?

2.70 Consider the situation of Exercise 2.69. Suppose 50,000 boxes of detergent are produced per week and suppose also that those underfilled are "sent back," with customers requesting reimbursement of the purchase price. Suppose also that the cost of production is known to be \$4.00 per box while the purchase price is \$4.50 per box.

- (a) What is the weekly profit under the condition of no defective boxes?
- (b) What is the loss in profit expected due to underfilling?

2.71 As the situation of Exercise 2.69 might suggest, statistical procedures are often used for control of quality (i.e., industrial quality control). At times, the weight of a product is an important variable to control. Specifications are given for the weight of a certain packaged product, and a package is rejected if it is either too light or too heavy. Historical data suggest that 0.95 is the probability that the product meets weight specifications whereas 0.002 is the probability that the product is too light. For each single packaged product, the manufacturer invests \$20.00 in production and the purchase price for the consumer is \$25.00.

- (a) What is the probability that a package chosen randomly from the production line is too heavy?
- (b) For each 10,000 packages sold, what profit is received by the manufacturer if all packages meet weight specification?
- (c) Assuming that all defective packages are rejected

and rendered worthless, how much is the profit reduced on 10,000 packages due to failure to meet weight specification?

2.72 Prove that

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

*See page 100.

2.6 Conditional Probability, Independence, and the Product Rule

One very important concept in probability theory is conditional probability. In some applications, the practitioner is interested in the probability structure under certain restrictions. For instance, in epidemiology, rather than studying the chance that a person from the general population has diabetes, it might be of more interest to know this probability for a distinct group such as Asian women in the age range of 35 to 50 or Hispanic men in the age range of 40 to 60. This type of probability is called a conditional probability.

Conditional Probability

The probability of an event B occurring when it is known that some event A has occurred is called a **conditional probability** and is denoted by $P(B|A)$. The symbol $P(B|A)$ is usually read “the probability that B occurs given that A occurs” or simply “the probability of B , given A .”

Consider the event B of getting a perfect square when a die is tossed. The die is constructed so that the even numbers are twice as likely to occur as the odd numbers. Based on the sample space $S = \{1, 2, 3, 4, 5, 6\}$, with probabilities of $1/9$ and $2/9$ assigned, respectively, to the odd and even numbers, the probability of B occurring is $1/3$. Now suppose that it is known that the toss of the die resulted in a number greater than 3. We are now dealing with a reduced sample space $A = \{4, 5, 6\}$, which is a subset of S . To find the probability that B occurs, relative to the space A , we must first assign new probabilities to the elements of A proportional to their original probabilities such that their sum is 1. Assigning a probability of w to the odd number in A and a probability of $2w$ to the two even numbers, we have $5w = 1$, or $w = 1/5$. Relative to the space A , we find that B contains the single element 4. Denoting this event by the symbol $B|A$, we write $B|A = \{4\}$, and hence

$$P(B|A) = \frac{2}{5}.$$

This example illustrates that events may have different probabilities when considered relative to different sample spaces.

We can also write

$$P(B|A) = \frac{2}{5} = \frac{2/9}{5/9} = \frac{P(A \cap B)}{P(A)},$$

where $P(A \cap B)$ and $P(A)$ are found from the original sample space S . In other words, a conditional probability relative to a subspace A of S may be calculated directly from the probabilities assigned to the elements of the original sample space S .

7 64 64

2.6 Conditional Probability, Independence, and the Product Rule

Definition 2.10: The conditional probability of B , given A , denoted by $P(B|A)$, is defined by

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad \text{provided } P(A) > 0.$$

As an additional illustration, suppose that our sample space S is the population of adults in a small town who have completed the requirements for a college degree. We shall categorize them according to gender and employment status. The data are given in Table 2.1.

Table 2.1: Categorization of the Adults in a Small Town

	Employed	Unemployed	Total
Male	460	40	500
Female	140	260	400
Total	600	300	900

One of these individuals is to be selected at random for a tour throughout the country to publicize the advantages of establishing new industries in the town. We shall be concerned with the following events:

M : a man is chosen,

E : the one chosen is employed.

Using the reduced sample space E , we find that

$$P(M|E) = \frac{460}{600} = \frac{23}{30}.$$

Let $n(A)$ denote the number of elements in any set A . Using this notation, since each adult has an equal chance of being selected, we can write

$$P(M|E) = \frac{n(E \cap M)}{n(E)} = \frac{n(E \cap M)/n(S)}{n(E)/n(S)} = \frac{P(E \cap M)}{P(E)},$$

where $P(E \cap M)$ and $P(E)$ are found from the original sample space S . To verify this result, note that

$$P(E) = \frac{600}{900} = \frac{2}{3} \quad \text{and} \quad P(E \cap M) = \frac{460}{900} = \frac{23}{45}.$$

Hence,

$$P(M|E) = \frac{23/45}{2/3} = \frac{23}{30},$$

as before.

Example 2.34: The probability that a regularly scheduled flight departs on time is $P(D) = 0.83$; the probability that it arrives on time is $P(A) = 0.82$; and the probability that it departs and arrives on time is $P(D \cap A) = 0.78$. Find the probability that a plane

(a) arrives on time, given that it departed on time, and (b) departed on time, given that it has arrived on time.

Solution: Using Definition 2.10, we have the following.

(a) The probability that a plane arrives on time, given that it departed on time, is

$$P(A|D) = \frac{P(D \cap A)}{P(D)} = \frac{0.78}{0.83} = 0.94.$$

(b) The probability that a plane departed on time, given that it has arrived on time, is

$$P(D|A) = \frac{P(D \cap A)}{P(A)} = \frac{0.78}{0.82} = 0.95.$$

The notion of conditional probability provides the capability of reevaluating the idea of probability of an event in light of additional information, that is, when it is known that another event has occurred. The probability $P(A|B)$ is an updating of $P(A)$ based on the knowledge that event B has occurred. In Example 2.34, it is important to know the probability that the flight arrives on time. One is given the information that the flight did not depart on time. Armed with this additional information, one can calculate the more pertinent probability $P(A|D')$, that is, the probability that it arrives on time, given that it did not depart on time. In many situations, the conclusions drawn from observing the more important conditional probability change the picture entirely. In this example, the computation of $P(A|D')$ is

$$P(A|D') = \frac{P(A \cap D')}{P(D')} = \frac{0.82 - 0.78}{0.17} = 0.24.$$

As a result, the probability of an on-time arrival is diminished severely in the presence of the additional information.

Example 2.35: The concept of conditional probability has countless uses in both industrial and biomedical applications. Consider an industrial process in the textile industry in which strips of a particular type of cloth are being produced. These strips can be defective in two ways, length and nature of texture. For the case of the latter, the process of identification is very complicated. It is known from historical information on the process that 10% of strips fail the length test, 5% fail the texture test, and only 0.8% fail both tests. If a strip is selected randomly from the process and a quick measurement identifies it as failing the length test, what is the probability that it is texture defective?

Solution: Consider the events

$$L: \text{length defective}, \quad T: \text{texture defective}.$$

Given that the strip is length defective, the probability that this strip is texture defective is given by

$$P(T|L) = \frac{P(T \cap L)}{P(L)} = \frac{0.008}{0.1} = 0.08.$$

Thus, knowing the conditional probability provides considerably more information than merely knowing $P(T)$.

Independent Events

In the die-tossing experiment discussed on page 62, we note that $P(B|A) = 2/5$ whereas $P(B) = 1/3$. That is, $P(B|A) \neq P(B)$, indicating that B depends on A . Now consider an experiment in which 2 cards are drawn in succession from an ordinary deck, with replacement. The events are defined as

A : the first card is an ace,

B : the second card is a spade.

Since the first card is replaced, our sample space for both the first and the second draw consists of 52 cards, containing 4 aces and 13 spades. Hence,

$$P(B|A) = \frac{13}{52} = \frac{1}{4} \quad \text{and} \quad P(B) = \frac{13}{52} = \frac{1}{4}.$$

That is, $P(B|A) = P(B)$. When this is true, the events A and B are said to be independent.

Although conditional probability allows for an alteration of the probability of an event in the light of additional material, it also enables us to understand better the very important concept of independence or, in the present context, independent events. In the airport illustration in Example 2.34, $P(A|D)$ differs from $P(A)$. This suggests that the occurrence of D influenced A , and this is certainly expected in this illustration. However, consider the situation where we have events A and B and

$$P(A|B) = P(A).$$

In other words, the occurrence of B had no impact on the odds of occurrence of A . Here the occurrence of A is independent of the occurrence of B . The importance of the concept of independence cannot be overemphasized. It plays a vital role in material in virtually all chapters in this book and in all areas of applied statistics.

Definition 2.11:

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

$$P(B|A) = P(B) \quad \text{or} \quad P(A|B) = P(A),$$

assuming the existences of the conditional probabilities. Otherwise, A and B are **dependent**.

The condition $P(B|A) = P(B)$ implies that $P(A|B) = P(A)$, and conversely. For the card-drawing experiments, where we showed that $P(B|A) = P(B) = 1/4$, we also can see that $P(A|B) = P(A) = 1/13$.

The Product Rule, or the Multiplicative Rule

Multiplying the formula in Definition 2.10 by $P(A)$, we obtain the following important **multiplicative rule** (or **product rule**), which enables us to calculate

the probability that two events will both occur.

Theorem 2.10:

If in an experiment the events A and B can both occur, then

$$P(A \cap B) = P(A)P(B|A), \text{ provided } P(A) > 0.$$

Thus, the probability that both A and B occur is equal to the probability that A occurs multiplied by the conditional probability that B occurs, given that A occurs. Since the events $A \cap B$ and $B \cap A$ are equivalent, it follows from Theorem 2.10 that we can also write

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

In other words, it does not matter which event is referred to as A and which event is referred to as B .

Example 2.36:

Suppose that we have a fuse box containing 20 fuses, of which 5 are defective. If 2 fuses are selected at random and removed from the box in succession without replacing the first, what is the probability that both fuses are defective?

Solution: We shall let A be the event that the first fuse is defective and B the event that the second fuse is defective; then we interpret $A \cap B$ as the event that A occurs and then B occurs after A has occurred. The probability of first removing a defective fuse is $1/4$; then the probability of removing a second defective fuse from the remaining 4 is $4/19$. Hence,

$$P(A \cap B) = \left(\frac{1}{4}\right)\left(\frac{4}{19}\right) = \frac{1}{19}.$$

Example 2.37:

One bag contains 4 white balls and 3 black balls, and a second bag contains 3 white balls and 5 black balls. One ball is drawn from the first bag and placed unseen in the second bag. What is the probability that a ball now drawn from the second bag is black?

Solution:

Let B_1 , B_2 , and W_1 represent, respectively, the drawing of a black ball from bag 1, a black ball from bag 2, and a white ball from bag 1. We are interested in the union of the mutually exclusive events $B_1 \cap B_2$ and $W_1 \cap B_2$. The various possibilities and their probabilities are illustrated in Figure 2.8. Now

$$\begin{aligned} P[(B_1 \cap B_2) \text{ or } (W_1 \cap B_2)] &= P(B_1 \cap B_2) + P(W_1 \cap B_2) \\ &= P(B_1)P(B_2|B_1) + P(W_1)P(B_2|W_1) \\ &= \left(\frac{3}{7}\right)\left(\frac{6}{9}\right) + \left(\frac{4}{7}\right)\left(\frac{5}{9}\right) = \frac{38}{63}. \end{aligned}$$

If, in Example 2.36, the first fuse is replaced and the fuses thoroughly rearranged before the second is removed, then the probability of a defective fuse on the second selection is still $1/4$; that is, $P(B|A) = P(B)$ and the events A and B are independent. When this is true, we can substitute $P(B)$ for $P(B|A)$ in Theorem 2.10 to obtain the following special multiplicative rule.

OR - Union
AND - Intersection

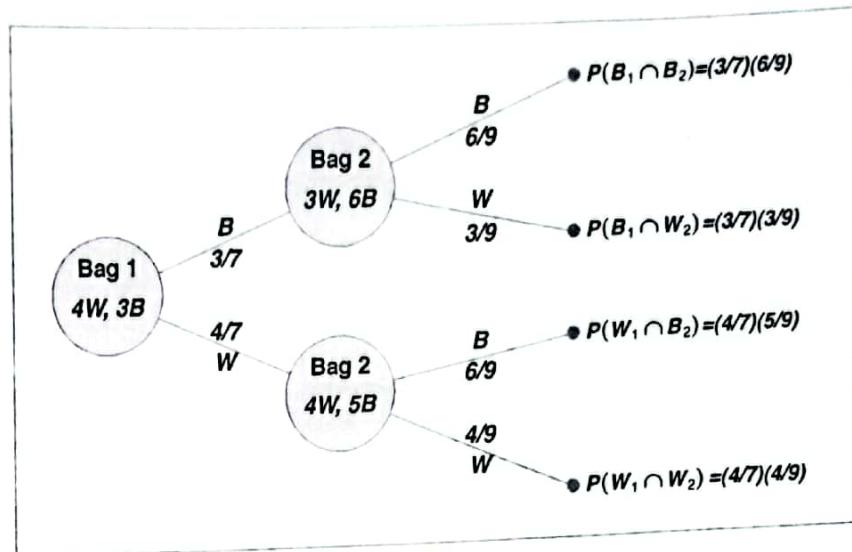


Figure 2.8: Tree diagram for Example 2.37.

Theorem 2.11: Two events A and B are independent if and only if

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

Therefore, to obtain the probability that two independent events will both occur, we simply find the product of their individual probabilities.

Example 2.38: A small town has one fire engine and one ambulance available for emergencies. The probability that the fire engine is available when needed is 0.98, and the probability that the ambulance is available when called is 0.92. In the event of an injury resulting from a burning building, find the probability that both the ambulance and the fire engine will be available, assuming they operate independently.

Solution: Let A and B represent the respective events that the fire engine and the ambulance are available. Then

Ans: i) $P(A \cap B) = P(A)P(B) = (0.98)(0.92) = 0.9016.$

Example 2.39: An electrical system consists of four components as illustrated in Figure 2.9. The system works if components A and B work and either of the components C or D works. The reliability (probability of working) of each component is also shown in Figure 2.9. Find the probability that (a) the entire system works and (b) the component C does not work, given that the entire system works. Assume that the four components work independently.

Solution: In this configuration of the system, A , B , and the subsystem C and D constitute a serial circuit system, whereas the subsystem C and D itself is a parallel circuit system.

(a) Clearly the probability that the entire system works can be calculated as

follows:

$$\begin{aligned} P[A \cap B \cap (C \cup D)] &= P(A)P(B)P(C \cup D) = P(A)P(B)[1 - P(C' \cap D')] \\ &= P(A)P(B)[1 - P(C')P(D')] \\ &= (0.9)(0.9)[1 - (1 - 0.8)(1 - 0.8)] = 0.7776. \end{aligned}$$

The equalities above hold because of the independence among the four components.

(b) To calculate the conditional probability in this case, notice that

$$\begin{aligned} P &= \frac{P(\text{the system works but } C \text{ does not work})}{P(\text{the system works})} \\ &= \frac{P(A \cap B \cap C' \cap D)}{P(\text{the system works})} = \frac{(0.9)(0.9)(1 - 0.8)(0.8)}{0.7776} = 0.1667. \end{aligned}$$

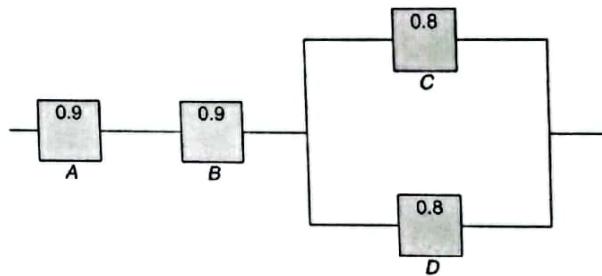


Figure 2.9: An electrical system for Example 2.39.

The multiplicative rule can be extended to more than two-event situations.

Theorem 2.12: If, in an experiment, the events A_1, A_2, \dots, A_k can occur, then

$$\begin{aligned} P(A_1 \cap A_2 \cap \dots \cap A_k) \\ = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \dots P(A_k|A_1 \cap A_2 \cap \dots \cap A_{k-1}). \end{aligned}$$

If the events A_1, A_2, \dots, A_k are independent, then

$$P(A_1 \cap A_2 \cap \dots \cap A_k) = P(A_1)P(A_2) \dots P(A_k).$$

Example 2.40: Three cards are drawn in succession, without replacement, from an ordinary deck of playing cards. Find the probability that the event $A_1 \cap A_2 \cap A_3$ occurs, where A_1 is the event that the first card is a red ace, A_2 is the event that the second card is a 10 or a jack, and A_3 is the event that the third card is greater than 3 but less than 7.

Solution: First we define the events

A_1 : the first card is a red ace;

A_2 : the second card is a 10 or a jack,

A_3 : the third card is greater than 3 but less than 7.

Now

$$P(A_1) = \frac{2}{52}, \quad P(A_2|A_1) = \frac{8}{51}, \quad P(A_3|A_1 \cap A_2) = \frac{12}{50},$$

and hence, by Theorem 2.12,

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \\ &= \left(\frac{2}{52}\right)\left(\frac{8}{51}\right)\left(\frac{12}{50}\right) = \frac{8}{5525}. \end{aligned}$$

The property of independence stated in Theorem 2.11 can be extended to deal with more than two events. Consider, for example, the case of three events A , B , and C . It is not sufficient to only have that $P(A \cap B \cap C) = P(A)P(B)P(C)$ as a definition of independence among the three. Suppose $A = B$ and $C = \emptyset$, the null set. Although $A \cap B \cap C = \emptyset$, which results in $P(A \cap B \cap C) = 0 = P(A)P(B)P(C)$, events A and B are not independent. Hence, we have the following definition.

Definition 2.12: A collection of events $\mathcal{A} = \{A_1, \dots, A_n\}$ are mutually independent if for any subset of \mathcal{A} , A_{i_1}, \dots, A_{i_k} , for $k \leq n$, we have

$$P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k}).$$

Exercises

2.73 If R is the event that a convict committed armed robbery and D is the event that the convict pushed dope, state in words what probabilities are expressed by

- (a) $P(R|D)$;
- (b) $P(D|R)$;
- (c) $P(R'|D')$.

2.74 A class in advanced physics is composed of 10 juniors, 30 seniors, and 10 graduate students. The final grades show that 3 of the juniors, 10 of the seniors, and 5 of the graduate students received an A for the course. If a student is chosen at random from this class and is found to have earned an A , what is the probability that he or she is a senior?

2.75 A random sample of 200 adults are classified below by sex and their level of education attained.

Education	Male	Female
Elementary	38	45
Secondary	28	50
College	22	17

If a person is picked at random from this group, find the probability that

- (a) the person is a male, given that the person has a secondary education;

- (b) the person does not have a college degree, given that the person is a female.

2.76 In an experiment to study the relationship of hypertension and smoking habits, the following data are collected for 180 individuals:

	Moderate Nonsmokers	Moderate Smokers	Heavy Smokers
H	21	36	30
NH	48	26	19

where H and NH in the table stand for *Hypertension* and *Nonhypertension*, respectively. If one of these individuals is selected at random, find the probability that the person is

- (a) experiencing hypertension, given that the person is a heavy smoker;
- (b) a nonsmoker, given that the person is experiencing no hypertension.

2.77 In the senior year of a high school graduating class of 100 students, 42 studied mathematics, 68 studied psychology, 54 studied history, 22 studied both mathematics and history, 25 studied both mathematics and psychology, 7 studied history but neither mathematics nor psychology, 10 studied all three subjects, and 8 did not take any of the three. Randomly select

a student from the class and find the probabilities of the following events.

- (a) A person enrolled in psychology takes all three subjects.
- (b) A person not taking psychology is taking both history and mathematics.

2.78 A manufacturer of a flu vaccine is concerned about the quality of its flu serum. Batches of serum are processed by three different departments having rejection rates of 0.10, 0.08, and 0.12, respectively. The inspections by the three departments are sequential and independent.

- (a) What is the probability that a batch of serum survives the first departmental inspection but is rejected by the second department?
- (b) What is the probability that a batch of serum is rejected by the third department?

2.79 In *USA Today* (Sept. 5, 1996), the results of a survey involving the use of sleepwear while traveling were listed as follows.

	Male	Female	Total
Underwear	0.220	0.024	0.244
Nightgown	0.002	0.180	0.182
Nothing	0.160	0.018	0.178
Pajamas	0.102	0.073	0.175
T-shirt	0.046	0.088	0.134
Other	0.084	0.003	0.087

- (a) What is the probability that a traveler is a female who sleeps in the nude?
- (b) What is the probability that a traveler is male?
- (c) Assuming the traveler is male, what is the probability that he sleeps in pajamas?
- (d) What is the probability that a traveler is male if the traveler sleeps in pajamas or a T-shirt?

2.80 The probability that an automobile being filled with gasoline also needs an oil change is 0.25; the probability that it needs a new oil filter is 0.40; and the probability that both the oil and the filter need changing is 0.14.

- (a) If the oil has to be changed, what is the probability that a new oil filter is needed?
- (b) If a new oil filter is needed, what is the probability that the oil has to be changed?

2.81 The probability that a married man watches a certain television show is 0.4, and the probability that a married woman watches the show is 0.5. The probability that a man watches the show, given that his wife does, is 0.7. Find the probability that

- (a) a married couple watches the show;

- (b) a wife watches the show, given that her husband does;
- (c) at least one member of a married couple will watch the show.

2.82 For married couples living in a certain suburb, the probability that the husband will vote on a bond referendum is 0.21, the probability that the wife will vote on the referendum is 0.28, and the probability that both the husband and the wife will vote is 0.15. What is the probability that

- (a) at least one member of a married couple will vote?
- (b) a wife will vote, given that her husband will vote?
- (c) a husband will vote, given that his wife will not vote?

2.83 The probability that a vehicle entering the Luray Caverns has Canadian license plates is 0.12; the probability that it is a camper is 0.28, and the probability that it is a camper with Canadian license plates is 0.09. What is the probability that

- (a) a camper entering the Luray Caverns has Canadian license plates?
- (b) a vehicle with Canadian license plates entering the Luray Caverns is a camper?
- (c) a vehicle entering the Luray Caverns does not have Canadian plates or is not a camper?

2.84 The probability that the head of a household is home when a telemarketing representative calls is 0.4. Given that the head of the house is home, the probability that goods will be bought from the company is 0.3. Find the probability that the head of the house is home and goods are bought from the company.

2.85 The probability that a doctor correctly diagnoses a particular illness is 0.7. Given that the doctor makes an incorrect diagnosis, the probability that the patient files a lawsuit is 0.9. What is the probability that the doctor makes an incorrect diagnosis and the patient sues?

2.86 In 1970, 11% of Americans completed four years of college; 43% of them were women. In 1990, 22% of Americans completed four years of college; 53% of them were women (*Time*, Jan 19, 1996).

- (a) Given that a person completed four years of college in 1970, what is the probability that the person was a woman?
- (b) What is the probability that a woman finished four years of college in 1990?
- (c) What is the probability that a man had not finished college in 1990?

2.87 A real estate agent has 8 master keys to open several new homes. Only 1 master key will open any given house. If 40% of these homes are usually left unlocked, what is the probability that the real estate agent can get into a specific home if the agent selects 3 master keys at random before leaving the office?

2.88 Before the distribution of certain statistical software, every fourth compact disk (CD) is tested for accuracy. The testing process consists of running four independent programs and checking the results. The failure rates for the four testing programs are, respectively, 0.01, 0.03, 0.02, and 0.01.

- What is the probability that a CD was tested and failed any test?
- Given that a CD was tested, what is the probability that it failed program 2 or 3?
- In a sample of 100, how many CDs would you expect to be rejected?
- Given that a CD was defective, what is the probability that it was tested?

2.89 A town has two fire engines operating independently. The probability that a specific engine is available when needed is 0.96.

- What is the probability that neither is available when needed?
- What is the probability that a fire engine is available when needed?

2.90 Pollution of the rivers in the United States has been a problem for many years. Consider the following events:

A : the river is polluted,

B : a sample of water tested detects pollution,

C : fishing is permitted.

Assume $P(A) = 0.3$, $P(B|A) = 0.75$, $P(B|A') = 0.20$, $P(C|A \cap B) = 0.20$, $P(C|A' \cap B) = 0.15$, $P(C|A \cap B') = 0.80$, and $P(C|A' \cap B') = 0.90$.

- Find $P(A \cap B \cap C)$
- Find $P(B' \cap C)$
- Find $P(C)$
- Find the probability that the river is polluted, given that fishing is permitted and the sample tested did not detect pollution.

2.91 Find the probability of randomly selecting 4 good quarts of milk in succession from a cooler containing 20 quarts of which 5 have spoiled, by using

- the first formula of Theorem 2.12 on page 68.
- the formulas of Theorem 2.6 and Rule 2.3 on pages 50 and 54, respectively.

2.92 Suppose the diagram of an electrical system is as given in Figure 2.10. What is the probability that the system works? Assume the components fail independently.

2.93 A circuit system is given in Figure 2.11. Assume the components fail independently.

- What is the probability that the entire system works?
- Given that the system works, what is the probability that the component A is not working?

2.94 In the situation of Exercise 2.93, it is known that the system does not work. What is the probability that the component A also does not work?

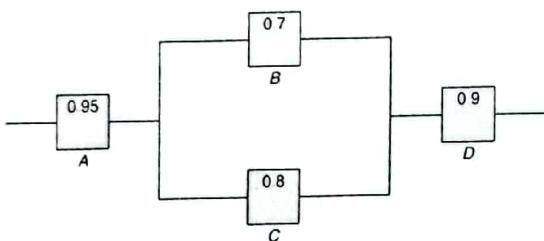


Figure 2.10: Diagram for Exercise 2.92.

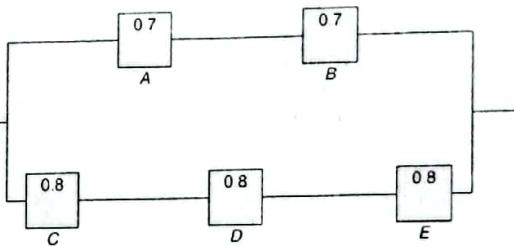


Figure 2.11: Diagram for Exercise 2.93.

2.7 Bayes' Rule

Bayesian statistics is a collection of tools that is used in a special form of statistical inference which applies in the analysis of experimental data in many practical situations in science and engineering. Bayes' rule is one of the most important rules in probability theory. It is the foundation of Bayesian inference, which will be discussed in Chapter 18.

Total Probability

Let us now return to the illustration of Section 2.6, where an individual is being selected at random from the adults of a small town to tour the country and publicize the advantages of establishing new industries in the town. Suppose that we are now given the additional information that 36 of those employed and 12 of those unemployed are members of the Rotary Club. We wish to find the probability of the event A that the individual selected is a member of the Rotary Club. Referring to Figure 2.12, we can write A as the union of the two mutually exclusive events $E \cap A$ and $E' \cap A$. Hence, $A = (E \cap A) \cup (E' \cap A)$, and by Corollary 2.1 of Theorem 2.7, and then Theorem 2.10, we can write

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

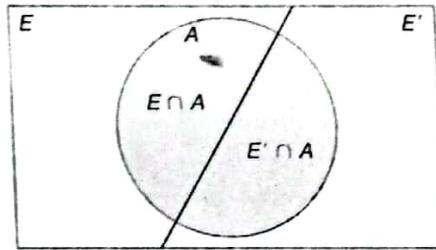


Figure 2.12: Venn diagram for the events A , E , and E' .

The data of Section 2.6, together with the additional data given above for the set A , enable us to compute

$$P(E) = \frac{600}{900} = \frac{2}{3}, \quad P(A|E) = \frac{36}{600} = \frac{3}{50},$$

and

$$P(E') = \frac{1}{3}, \quad P(A|E') = \frac{12}{300} = \frac{1}{25}.$$

If we display these probabilities by means of the tree diagram of Figure 2.13, where the first branch yields the probability $P(E)P(A|E)$ and the second branch yields

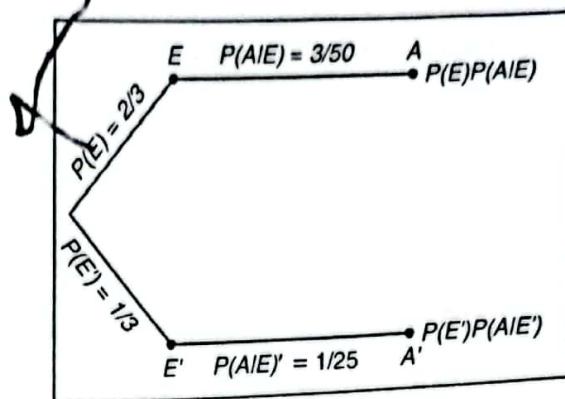


Figure 2.13: Tree diagram for the data on page 63, using additional information on page 72.

the probability $P(E')P(A|E')$, it follows that

$$P(A) = \left(\frac{2}{3}\right)\left(\frac{3}{50}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{25}\right) = \frac{4}{75}.$$

A generalization of the foregoing illustration to the case where the sample space is partitioned into k subsets is covered by the following theorem, sometimes called the **theorem of total probability** or the **rule of elimination**.

Theorem 2.13:

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

Total probability

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

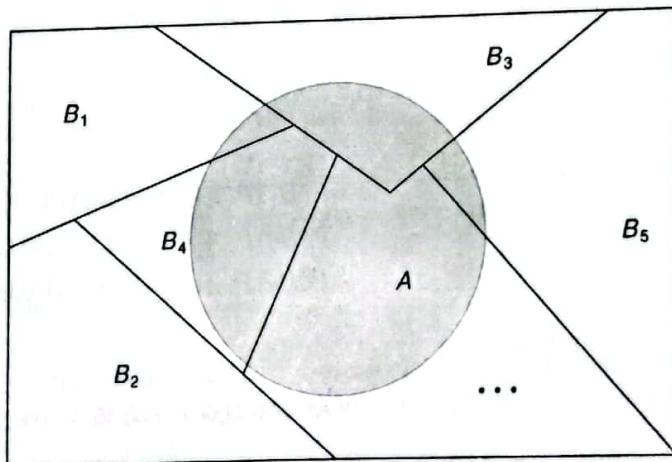


Figure 2.14: Partitioning the sample space S .

Proof: Consider the Venn diagram of Figure 2.14. The event A is seen to be the union of the mutually exclusive events

$$B_1 \cap A, B_2 \cap A, \dots, B_k \cap A;$$

that is,

$$A = (B_1 \cap A) \cup (B_2 \cap A) \cup \dots \cup (B_k \cap A).$$

Using Corollary 2.2 of Theorem 2.7 and Theorem 2.10, we have

$$\begin{aligned} P(A) &= P[(B_1 \cap A) \cup (B_2 \cap A) \cup \dots \cup (B_k \cap A)] \\ &= P(B_1 \cap A) + P(B_2 \cap A) + \dots + P(B_k \cap A) \\ &= \sum_{i=1}^k P(B_i \cap A) \\ &= \sum_{i=1}^k P(B_i)P(A|B_i). \end{aligned}$$

Example 2.41: In a certain assembly plant, three machines, B_1 , B_2 , and B_3 , make 30%, 45%, and 25%, respectively, of the products. It is known from past experience that 2%, 3%, and 2% of the products made by each machine, respectively, are defective. Now, suppose that a finished product is randomly selected. What is the probability that it is defective?

Solution: Consider the following events:

A : the product is defective,

E_1 : B_1 : the product is made by machine B_1 ,

E_2 : the product is made by machine B_2 ,

E_3 : the product is made by machine B_3 .

Applying the rule of elimination, we can write

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3).$$

Referring to the tree diagram of Figure 2.15, we find that the three branches give the probabilities

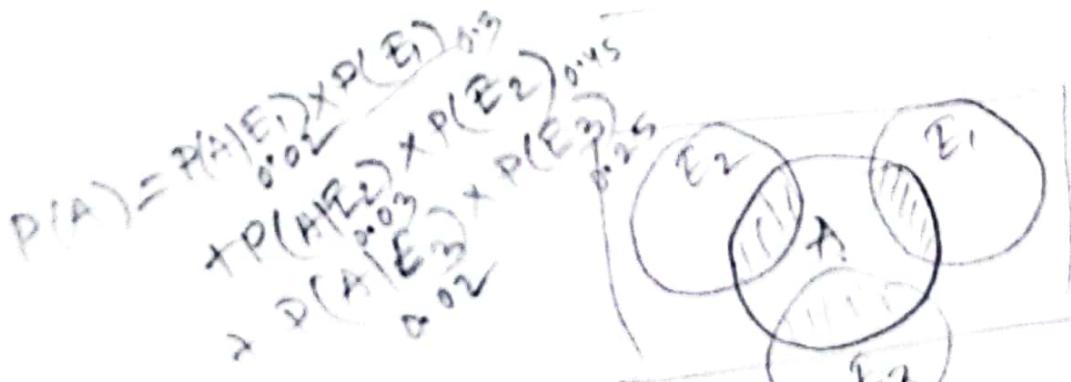
$$P(B_1)P(A|B_1) = (0.3)(0.02) = 0.006,$$

$$P(B_2)P(A|B_2) = (0.45)(0.03) = 0.0135,$$

$$P(B_3)P(A|B_3) = (0.25)(0.02) = 0.005,$$

and hence

$$P(A) = 0.006 + 0.0135 + 0.005 = 0.0245.$$



$P(A) = P(A|E_1)P(E_1) + P(A|E_2)P(E_2) + P(A|E_3)P(E_3)$

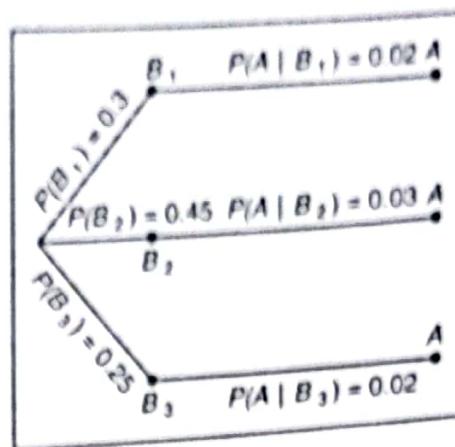


Figure 2.15: Tree diagram for Example 2.41.

Bayes' Rule

Instead of asking for $P(A)$ in Example 2.41, by the rule of elimination, suppose that we now consider the problem of finding the conditional probability $P(B_i|A)$. In other words, suppose that a product was randomly selected and it is defective. What is the probability that this product was made by machine B_i ? Questions of this type can be answered by using the following theorem, called **Bayes' rule**:

Theorem 2.14: **(Bayes' Rule)** If the events B_1, B_2, \dots, B_k constitute a partition of the sample space S such that $P(B_i) \neq 0$ for $i = 1, 2, \dots, k$, then for any event A in S such that $P(A) \neq 0$,

$$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^k P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^k P(B_i)P(A|B_i)} \quad \text{for } r = 1, 2, \dots, k.$$

Proof: By the definition of conditional probability,

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

and then using Theorem 2.13 in the denominator, we have

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

which completes the proof. ■

Example 2.42: With reference to Example 2.41, if a product was chosen randomly and found to be defective, what is the probability that it was made by machine B_3 ?

Solution: Using Bayes' rule to write

$$P(B_3|A) = \frac{P(B_3)P(A|B_3)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)},$$

$$P(B_3|A) = \frac{P(B_3 \cap A)}{P(A)}$$

and then substituting the probabilities calculated in Example 2.41, we have

$$P(B_3|A) = \frac{0.005}{0.006 + 0.0135 + 0.005} = \frac{0.005}{0.0245} = \frac{10}{49}.$$

In view of the fact that a defective product was selected, this result suggests that it probably was not made by machine B_3 . ■

Example 2.43: *H.W.* A manufacturing firm employs three analytical plans for the design and development of a particular product. For cost reasons, all three are used at varying times. In fact, plans 1, 2, and 3 are used for 30%, 20%, and 50% of the products, respectively. The defect rate is different for the three procedures as follows:

$$P(D|P_1) = 0.01, \quad P(D|P_2) = 0.03, \quad P(D|P_3) = 0.02,$$

where $P(D|P_j)$ is the probability of a defective product, given plan j . If a random product was observed and found to be defective, which plan was most likely used and thus responsible?

Solution: From the statement of the problem

$$P(P_1) = 0.30, \quad P(P_2) = 0.20, \quad \text{and} \quad P(P_3) = 0.50,$$

we must find $P(P_j|D)$ for $j = 1, 2, 3$. Bayes' rule (Theorem 2.14) shows

$$\begin{aligned} P(P_1|D) &= \frac{P(P_1)P(D|P_1)}{P(P_1)P(D|P_1) + P(P_2)P(D|P_2) + P(P_3)P(D|P_3)} \\ &= \frac{(0.30)(0.01)}{(0.3)(0.01) + (0.20)(0.03) + (0.50)(0.02)} = \frac{0.003}{0.019} = 0.158. \end{aligned}$$

Similarly,

$$P(P_2|D) = \frac{(0.03)(0.20)}{0.019} = 0.316 \quad \text{and} \quad P(P_3|D) = \frac{(0.02)(0.50)}{0.019} = 0.526.$$

The conditional probability of a defect given plan 3 is the largest of the three; thus a defective for a random product is most likely the result of the use of plan 3. ■

Using Bayes' rule, a statistical methodology called the Bayesian approach has attracted a lot of attention in applications. An introduction to the Bayesian method will be discussed in Chapter 18.

Exercises

- 2.95** In a certain region of the country it is known from past experience that the probability of selecting an adult over 40 years of age with cancer is 0.05. If the probability of a doctor correctly diagnosing a person with cancer as having the disease is 0.78 and the probability of incorrectly diagnosing a person without cancer as having the disease is 0.06, what is the prob-

ability that an adult over 40 years of age is diagnosed as having cancer?

- 2.96** Police plan to enforce speed limits by using radar traps at four different locations within the city limits. The radar traps at each of the locations L_1, L_2, L_3 , and L_4 will be operated 40%, 30%, 20%, and 30% of

3582289951018 *151*
6818158809597

time. If a person who is speeding on her way to work has probabilities of 0.2, 0.1, 0.5, and 0.2, respectively, of passing through these locations, what is the probability that she will receive a speeding ticket?

2.97 Referring to Exercise 2.95, what is the probability that a person diagnosed as having cancer actually has the disease?

2.98 If the person in Exercise 2.96 received a speeding ticket on her way to work, what is the probability that she passed through the radar trap located at L_2 ?

2.99 Suppose that the four inspectors at a film factory are supposed to stamp the expiration date on each package of film at the end of the assembly line. John, who stamps 20% of the packages, fails to stamp the expiration date once in every 200 packages. Tom, who stamps 60% of the packages, fails to stamp the expiration date once in every 100 packages. Jeff, who stamps 15% of the packages, fails to stamp the expiration date once in every 90 packages, and Pat, who stamps 5% of the packages, fails to stamp the expiration date once in every 200 packages. If a customer complains that her package of film does not show the expiration date, what is the probability that it was inspected by John?

2.100 A regional telephone company operates three identical relay stations at different locations. During a

one-year period, the number of malfunctions reported by each station and the causes are shown below.

Station	A	B	C
Problems with electricity supplied	2	1	1
Computer malfunction	4	3	2
Malfunctioning electrical equipment	5	4	2
Caused by other human errors	7	7	5

Suppose that a malfunction was reported and it was found to be caused by other human errors. What is the probability that it came from station C?

2.101 A paint-store chain produces and sells latex and semigloss paint. Based on long-range sales, the probability that a customer will purchase latex paint is 0.75. Of those that purchase latex paint, 60% also purchase rollers. But only 30% of semigloss paint buyers purchase rollers. A randomly selected buyer purchases a roller and a can of paint. What is the probability that the paint is latex?

2.102 Denote by A , B , and C the events that a grand prize is behind doors A , B , and C , respectively. Suppose you randomly picked a door, say A . The game host opened a door, say B , and showed there was no prize behind it. Now the host offers you the option of either staying at the door that you picked (A) or switching to the remaining unopened door (C). Use probability to explain whether you should switch or not.

Review Exercises

2.103 A truth serum has the property that 90% of the guilty suspects are properly judged while, of course, 10% of the guilty suspects are improperly found innocent. On the other hand, innocent suspects are misjudged 1% of the time. If the suspect was selected from a group of suspects of which only 5% have ever committed a crime, and the serum indicates that he is guilty, what is the probability that he is innocent?

2.104 An allergist claims that 50% of the patients she tests are allergic to some type of weed. What is the probability that

- (a) exactly 3 of her next 4 patients are allergic to weeds?
- (b) none of her next 4 patients is allergic to weeds?

2.105 By comparing appropriate regions of Venn diagrams, verify that

- (a) $(A \cap B) \cup (A \cap B') = A$;
- (b) $A' \cap (B' \cup C) = (A' \cap B') \cup (A' \cap C)$.

2.106 The probabilities that a service station will pump gas into 0, 1, 2, 3, 4, or 5 or more cars during a certain 30-minute period are 0.03, 0.18, 0.24, 0.28, 0.10, and 0.17, respectively. Find the probability that in this 30-minute period

- (a) more than 2 cars receive gas;
- (b) at most 4 cars receive gas;
- (c) 4 or more cars receive gas.

2.107 How many bridge hands are possible containing 4 spades, 6 diamonds, 1 club, and 2 hearts?

2.108 If the probability is 0.1 that a person will make a mistake on his or her state income tax return, find the probability that

- (a) four totally unrelated persons each make a mistake;
- (b) Mr. Jones and Ms. Clark both make mistakes, and Mr. Roberts and Ms. Williams do not make a mistake.

2.109 A large industrial firm uses three local motels to provide overnight accommodations for its clients. From past experience it is known that 20% of the clients are assigned rooms at the Ramada Inn, 50% at the Sheraton, and 30% at the Lakeview Motor Lodge. If the plumbing is faulty in 5% of the rooms at the Ramada Inn, in 4% of the rooms at the Sheraton, and in 8% of the rooms at the Lakeview Motor Lodge, what is the probability that

- (a) a client will be assigned a room with faulty plumbing?
- (b) a person with a room having faulty plumbing was assigned accommodations at the Lakeview Motor Lodge?

2.110 The probability that a patient recovers from a delicate heart operation is 0.8. What is the probability that

- (a) exactly 2 of the next 3 patients who have this operation survive?
- (b) all of the next 3 patients who have this operation survive?

2.111 In a certain federal prison, it is known that $\frac{2}{3}$ of the inmates are under 25 years of age. It is also known that $\frac{3}{5}$ of the inmates are male and that $\frac{5}{8}$ of the inmates are female or 25 years of age or older. What is the probability that a prisoner selected at random from this prison is female and at least 25 years old?

2.112 From 4 red, 5 green, and 6 yellow apples, how many selections of 9 apples are possible if 3 of each color are to be selected?

2.113 From a box containing 6 black balls and 4 green balls, 3 balls are drawn in succession, each ball being replaced in the box before the next draw is made. What is the probability that

- (a) all 3 are the same color?
- (b) each color is represented?

2.114 A shipment of 12 television sets contains 3 defective sets. In how many ways can a hotel purchase 5 of these sets and receive at least 2 of the defective sets?

2.115 A certain federal agency employs three consulting firms (*A*, *B*, and *C*) with probabilities 0.40, 0.35, and 0.25, respectively. From past experience it is known that the probability of cost overruns for the firms are 0.05, 0.03, and 0.15, respectively. Suppose a cost overrun is experienced by the agency.

- (a) What is the probability that the consulting firm involved is company *C*?
- (b) What is the probability that it is company *A*?

2.116 A manufacturer is studying the effects of cooking temperature, cooking time, and type of cooking oil for making potato chips. Three different temperatures, 4 different cooking times, and 3 different oils are to be used

- (a) What is the total number of combinations to be studied?
- (b) How many combinations will be used for each type of oil?
- (c) Discuss why permutations are not an issue in this exercise

2.117 Consider the situation in Exercise 2.116, and suppose that the manufacturer can try only two combinations in a day.

- (a) What is the probability that any given set of two runs is chosen?
- (b) What is the probability that the highest temperature is used in either of these two combinations?

2.118 A certain form of cancer is known to be found in women over 60 with probability 0.07. A blood test exists for the detection of the disease, but the test is not infallible. In fact, it is known that 10% of the time the test gives a false negative (i.e., the test incorrectly gives a negative result) and 5% of the time the test gives a false positive (i.e., incorrectly gives a positive result). If a woman over 60 is known to have taken the test and received a favorable (i.e., negative) result, what is the probability that she has the disease?

2.119 A producer of a certain type of electronic component ships to suppliers in lots of twenty. Suppose that 60% of all such lots contain no defective components, 30% contain one defective component, and 10% contain two defective components. A lot is picked, two components from the lot are randomly selected and tested, and neither is defective.

- (a) What is the probability that zero defective components exist in the lot?
- (b) What is the probability that one defective exists in the lot?
- (c) What is the probability that two defectives exist in the lot?

2.120 A rare disease exists with which only 1 in 500 is affected. A test for the disease exists, but of course it is not infallible. A correct positive result (patient actually has the disease) occurs 95% of the time, while a false positive result (patient does not have the dis-

(case) occurs 1% of the time. If a randomly selected individual is tested and the result is positive, what is the probability that the individual has the disease?

2.121 A construction company employs two sales engineers. Engineer 1 does the work of estimating cost for 70% of jobs bid by the company. Engineer 2 does the work for 30% of jobs bid by the company. It is known that the error rate for engineer 1 is such that 0.02 is the probability of an error when he does the work, whereas the probability of an error in the work of engineer 2 is 0.04. Suppose a bid arrives and a serious error occurs in estimating cost. Which engineer would you guess did the work? Explain and show all work.

2.122 In the field of quality control, the science of statistics is often used to determine if a process is "out of control." Suppose the process is, indeed, out of control and 20% of items produced are defective.

- (a) If three items arrive off the process line in succession, what is the probability that all three are defective?
- (b) If four items arrive in succession, what is the probability that three are defective?

2.123 An industrial plant is conducting a study to determine how quickly injured workers are back on the job following injury. Records show that 10% of all injured workers are admitted to the hospital for treatment and 15% are back on the job the next day. In addition, studies show that 2% are both admitted for hospital treatment and back on the job the next day. If a worker is injured, what is the probability that the worker will either be admitted to a hospital or be back on the job the next day or both?

2.124 A firm is accustomed to training operators who do certain tasks on a production line. Those operators who attend the training course are known to be able to meet their production quotas 90% of the time. New operators who do not take the training course only meet their quotas 65% of the time. Fifty percent of new operators attend the course. Given that a new operator meets her production quota, what is the probability that she attended the program?

2.125 A survey of those using a particular statistical software system indicated that 10% were dissatisfied.

Half of those dissatisfied purchased the system from vendor A. It is also known that 20% of those surveyed purchased from vendor A. Given that the software was purchased from vendor A, what is the probability that that particular user is dissatisfied?

2.126 During bad economic times, industrial workers are dismissed and are often replaced by machines. The history of 100 workers whose loss of employment is attributable to technological advances is reviewed. For each of these individuals it is determined if he or she was given an alternative job within the same company, found a job with another company in the same field, found a job in a new field, or has been unemployed for 1 year. In addition, the union status of each worker is recorded. The following table summarizes the results

	Union	Nonunion
Same Company	40	15
New Company (same field)	13	10
New Field	4	11
Unemployed	2	5

- (a) If the selected worker found a job with a new company in the same field, what is the probability that the worker is a union member?
- (b) If the worker is a union member, what is the probability that the worker has been unemployed for a year?

2.127 There is a 50-50 chance that the queen carries the gene of hemophilia. If she is a carrier, then each prince has a 50-50 chance of having hemophilia independently. If the queen is not a carrier, the prince will not have the disease. Suppose the queen has had three princes without the disease. What is the probability the queen is a carrier?

2.128 Group Project: Give each student a bag of chocolate M&Ms. Divide the students into groups of 5 or 6. Calculate the relative frequency distribution for color of M&Ms for each group.

- (a) What is your estimated probability of randomly picking a yellow? a red?
- (b) Redo the calculations for the whole classroom. Did the estimates change?
- (c) Do you believe there is an equal number of each color in a process batch? Discuss.

2.8 Potential Misconceptions and Hazards: Relationship to Material in Other Chapters

This chapter contains the fundamental definitions, rules, and theorems that provide a foundation that renders probability an important tool for evaluating

scientific and engineering systems. The evaluations are often in the form of probability computations, as is illustrated in examples and exercises. Concepts such as independence, conditional probability, Bayes' rule, and others tend to mesh nicely to solve practical problems in which the bottom line is to produce a probability value. Illustrations in exercises are abundant. See, for example, Exercises 2.100 and 2.101. In these and many other exercises, an evaluation of a scientific system is being made judiciously from a probability calculation, using rules and definitions discussed in the chapter.

Now, how does the material in this chapter relate to that in other chapters? It is best to answer this question by looking ahead to Chapter 3. Chapter 3 also deals with the type of problems in which it is important to calculate probabilities. We illustrate how system performance depends on the value of one or more probabilities. Once again, conditional probability and independence play a role. However, new concepts arise which allow more structure based on the notion of a random variable and its probability distribution. Recall that the idea of frequency distributions was discussed briefly in Chapter 1. The probability distribution displays, in equation form or graphically, the total information necessary to describe a probability structure. For example, in Review Exercise 2.122 the random variable of interest is the number of defective items, a discrete measurement. Thus, the probability distribution would reveal the probability structure for the number of defective items out of the number selected from the process. As the reader moves into Chapter 3 and beyond, it will become apparent that assumptions will be required in order to determine and thus make use of probability distributions for solving scientific problems.