Basic Probability Concepts

1.1 Introduction

Probability deals with unpredictability and randomness, and probability theory is the branch of mathematics that is concerned with the study of random phenomena. A random phenomenon is one that, under repeated observation, yields different outcomes that are not deterministically predictable. However, these outcomes obey certain conditions of statistical regularity whereby the relative frequency of occurrence of the possible outcomes is approximately predictable. Examples of these random phenomena include the number of electronic mail (e-mail) messages received by all employees of a company in one day, the number of phone calls arriving at the university's switchboard over a given period, the number of components of a system that fail within a given interval, and the number of A's that a student can receive in one academic year.

According to the preceding definition, the fundamental issue in random phenomena is the idea of a repeated experiment with a set of possible outcomes or events. Associated with each of these events is a real number called the probability of the event that is related to the frequency of occurrence of the event in a long sequence of repeated trials of the experiment. In this way it becomes obvious that the probability of an event is a value that lies between zero and one, and the sum of the probabilities of the events for a particular experiment should sum to one.

This chapter begins with events associated with a random experiment. Then it provides different definitions of probability and considers elementary set theory and algebra of sets. Finally, it discusses basic concepts in combinatorial analysis that will be used in many of the later chapters.

1.2 Sample Space and Events

The concepts of *experiments* and *events* are very important in the study of probability. In probability, an experiment is any process of trial and observation. An experiment whose outcome is uncertain before it is performed is called a *random* experiment. When we perform a random experiment, the collection of possible elementary outcomes is called the *sample space* of the experiment, which is usually denoted by S. We define these outcomes as elementary outcomes because exactly one of the outcomes occurs when the experiment is performed. The elementary outcomes of an experiment are called the *sample points* of the sample space and are denoted by $w_i, i = 1, 2, \ldots$. If there are n possible outcomes of an experiment, then the sample space is $S = \{w_1, w_2, \ldots, w_n\}$.

An *event* is the occurrence of either a prescribed outcome or any one of a number of possible outcomes of an experiment. Thus, an event is a subset of the sample space. For example, if we toss a die, any number from 1 to 6 may appear. Therefore, in this experiment the sample space is defined by

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

The event "the outcome of the toss of a die is an even number" is a subset of S and is defined by

$$E = \{2, 4, 6\}$$

For a second example, consider a coin-tossing experiment in which each toss can result in either a head (H) or a tail (T). If we toss a coin three times and let the triplet xyz denote the outcome "x on the first toss, y on the second toss, and z on the third toss," then the sample space of the experiment is

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

The event "one head and two tails" is a subset of S and is defined by

$$E = \{HTT, THT, TTH\}$$

Other examples of events are as follows:

- In a single coin toss experiment with sample space $S = \{H, T\}$, the event $E = \{H\}$ is the event that a head appears on the toss and $E = \{T\}$ is the event that a tail appears on the toss.
- If we toss a coin twice and let xy denote the outcome "x on the first toss and y on the second toss," where x is head or tail and y is head or tail, then the sample space is $S = \{HH, HT, TH, TT\}$. The event $E = \{HT, TT\}$ is the event that a tail appears on the second toss.
- If we measure the lifetime of an electronic component, such as a chip, the sample space consists of all nonnegative real numbers. That is,

  $$S = \{x | 0 \leq x < \infty\}$$

  The event that the lifetime is not more than 7 hours is defined as follows:

  $$E = \{x | 0 \leq x \leq 7\}$$

- If we toss a die twice and let the pair (x, y) denote the outcome "x on the first toss and y on the second toss," then the sample space is

  $$S = \left\{ \begin{array}{cccccc} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{array} \right\}$$

  The event that the sum of the two tosses is 8 is denoted by

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

For any two events A and B defined on a sample space S, we can define the following new events:

- $A \cup B$ is the event that consists of all sample points that are either in A or in B or in both A and B. The event $A \cup B$ is called the *union* of events A and B.
- $A \cap B$ is the event that consists of all sample points that are in both A and B. The event $A \cap B$ is called the *intersection* of events A and B. Two events are

defined to be *mutually exclusive* if their intersection does not contain a sample point; that is, they have no outcomes in common. Events $A_1, A_2, A_3, \ldots,$ are defined to be mutually exclusive if no two of them have any outcomes in common and the events collectively have no outcomes in common.

- $A - B$ is the event that consists of all sample points that are in A but not in B. The event $A - B$ is called the *difference* of events A and B. Note that $A - B$ is different from $B - A$.

The algebra of unions, intersections, and differences of events will be discussed in greater detail when we study set theory later in this chapter.

1.3 Definitions of Probability

There are several ways to define probability. In this section we consider three definitions: the *axiomatic* definition, the *relative-frequency* definition, and the *classical* definition.

1.3.1 Axiomatic Definition

Consider a random experiment whose sample space is S. For each event A of S we assume that a number $P(A)$, called the *probability* of event A, is defined such that the following hold:

1. **Axiom 1:** $0 \leq P(A) \leq 1$, which means that the probability of A is some number between and including 0 and 1.
2. **Axiom 2:** $P(S) = 1$, which states that with probability 1, the outcome will be a sample point in the sample space.
3. **Axiom 3:** For any set of n mutually exclusive events $A_1, A_2, \ldots, A_n$ defined on the same sample space,

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

   That is, for any set of mutually exclusive events defined on the same space, the probability of at least one of these events occurring is the sum of their respective probabilities.

1.3.2 Relative-Frequency Definition

Consider a random experiment that is performed n times. If an event A occurs n_A times, then the probability of event A, $P(A)$, is defined as follows:

$$P(A) = \lim_{n \to \infty} \frac{n_A}{n}$$

The ratio n_A/n is called the *relative frequency* of event A. While the relative-frequency definition of probability is intuitively satisfactory for many practical problems, it has a few limitations. One such limitation is the fact that the experiment may not be repeatable, especially when we are dealing with destructive testing of expensive and/or scarce resources. Also, the limit may not exist.

1.3.3 Classical Definition

In the classical definition, the probability $P(A)$ of an event A is the ratio of the number of outcomes N_{A} of an experiment that are favorable to A to the total number N of possible outcomes of the experiment. That is,

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

This probability is determined *a priori* without actually performing the experiment. For example, in a coin toss experiment, there are two possible outcomes: heads or tails. Thus, $N = 2$, and if the coin is fair, the probability of the event that the toss comes up heads is $1/2$.

Example 1.1 Two fair dice are tossed. Find the probability of each of the following events:

a. The sum of the outcomes of the two dice is equal to 7
b. The sum of the outcomes of the two dice is equal to 7 or 11
c. The outcome of the second die is greater than the outcome of the first die
d. Both dice come up with even numbers

Solution We first define the sample space of the experiment. If we let the pair (x, y) denote the outcome "first die comes up x and second die comes up y," where $x, y \in \{1, 2, 3, 4, 5, 6\}$, then $S = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6), (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6), (5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6), (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)\}$. The total number of sample points is 36. We evaluate the three probabilities using the classical definition method.

(a) Let A_1 denote the event that the sum of the outcomes of the two dice is equal to seven. Then $A_1 = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$. Since the number of sample points in the event is 6, we have that $P(A_1) = 6/36 = 1/6$.

(b) Let B denote the event that the sum of the outcomes of the two dice is either seven or eleven, and let A_2 denote the event that the sum of the outcomes of the two dice is eleven. Then, $A_2 = \{(5,6),(6,5)\}$ with 2 sample points. Thus, $P(A_2) = 2/36 = 1/18$. Since B is the union of A_1 and A_2, which are mutually exclusive events, we obtain

$$P(B) = P(A_1 \cup A_2) = P(A_1) + P(A_2) = \frac{1}{6} + \frac{1}{18} = \frac{2}{9}$$

(c) Let C denote the event that the outcome of the second die is greater than the outcome of the first die. Then $C = \{(1,2),(1,3),(1,4),(1,5),(1,6),(2,3),(2,4),(2,5),(2,6),(3,4),(3,5),(3,6),(4,5),(4,6),(5,6)\}$ with 15 sample points. Thus, $P(C) = 15/36 = 5/12$.

(d) Let D denote the event that both dice come up with even numbers. Then $D = \{(2,2),(2,4),(2,6),(4,2),(4,4),(4,6),(6,2),(6,4),(6,6)\}$ with 9 sample points. Thus, $P(D) = 9/36 = 1/4$.

Note that the problem can also be solved by considering a two-dimensional display of the sample space, as shown in Figure 1.1. The figure shows the different events just defined.

▲

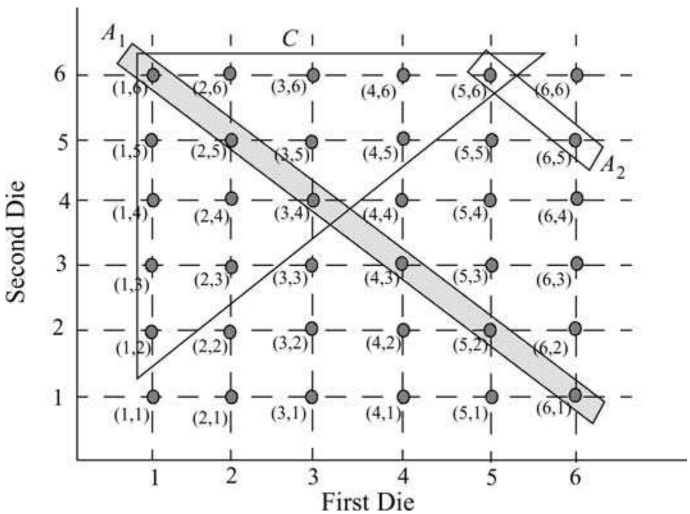

Figure 1.1 Sample Space for Example 1.1

The sample points in event D are spread over the entire sample space. Therefore, the event D is not shown in Figure 1.1.

1.4 Applications of Probability

There are several science and engineering applications of probability. Some of these applications are as follows.

1.4.1 Reliability Engineering

Reliability theory is concerned with the duration of the useful life of components and systems of components. System failure times are unpredictable. Thus, the time until a system fails, which is referred to as the *time to failure* of the system, is usually modeled by a probabilistic function. Reliability applications of probability are considered later in this chapter.

1.4.2 Quality Control

Quality control deals with the inspection of finished products to ensure that they meet the desired requirements and specifications. One way to perform the quality control function is to physically test/inspect each product as it comes off the production line. However, this is a very costly way to do it. The practical method is to randomly select a sample of the product from a lot and test each item in the sample. A decision to declare the lot good or defective is thus based on the outcome of the test of the items of the sample. This decision is itself based on a well-designed policy that guarantees that a good lot is rejected with a very small probability and that a bad lot is accepted with a very small probability. A lot is considered good if the parameter that characterizes the quality of the sample has a value that exceeds a predefined threshold value. Similarly the lot is considered to be defective if the parameter that characterizes the quality of the sample has a value that is smaller than the predefined threshold value. For example, one rule for acceptance of a lot can be that the number of defective items in the selected sample be less than some predefined fraction of the sample; otherwise the lot is declared defective.

1.4.3 Channel Noise

Noise is an unwanted signal. A message transmitted from a source passes through a channel where it is subject to different kinds of random disturbances that can introduce errors in the message received at the sink. That is, channel noise corrupts messages, as shown in Figure 1.2.

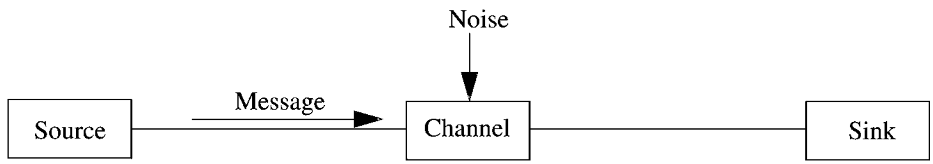

Figure 1.2 Model of a Communication System

Since noise is a random signal, one of the performance issues is the probability that the received message was not corrupted by noise. Thus, probability plays an important role in evaluating the performance of noisy communication channels.

1.4.4 System Simulation

Sometimes it is difficult to provide an exact solution of physical problems involving random phenomena. The difficulty arises from the fact that such problems are very complex, which is the case, for example, when a system has unusual properties. One way to deal with these problems is to provide an approximate solution, which attempts to make simplifying assumptions that enable the problem to be solved analytically. Another method is to use computer simulation, which imitates the physical process. Even when an approximate solution is obtained, it is always advisable to use simulation to validate the assumptions.

A simulation model describes the operation of a system in terms of individual events of the individual elements in the system. The model includes the interrelationships among the different elements and allows the effects of the elements' actions on each other to be captured as a dynamic process.

The key to a simulation model is the generation of random numbers that can be used to represent events—such as arrival of customers at a bank—in the system being modeled. Because these events are random in nature, the random numbers are used to drive the probability distributions that characterize them. Thus, knowledge of probability theory is essential for a meaningful simulation analysis.

1.5 Elementary Set Theory

A set is a collection of objects known as elements. The events that we discussed earlier in this chapter are usually modeled as sets, and the algebra of sets is used to study events. A set can be represented in a number of ways as the following examples illustrate.

Let A denote the set of positive integers between and including 1 and 5. Then

$$A = \{a | 1 \leq a \leq 5\} = \{1, 2, 3, 4, 5\}$$

Similarly, let B denote the set of positive odd numbers less than 10. Then

$$B = \{1, 3, 5, 7, 9\}$$

If k is an element of the set E, we say that k belongs to (or is a member of) E and write $k \in E$. If k is not an element of the set E, we say that k does not belong to (or is not a member of) E and write $k \notin E$.

A set A is called a *subset* of set B, denoted by $A \subset B$, if every member of A is a member of B. Alternatively, we say that the set B contains the set A by writing $B \supset A$.

The set that contains all possible elements is called the *universal set* S. The set that contains no elements (or is empty) is called the *null set* $\varnothing$ (or *empty set*).

1.5.1 Set Operations

Equality. Two sets A and B are defined to be equal, denoted by $A = B$, if and only if (iff) A is a subset of B and B is a subset of A; that is $A \subset B$, and $B \subset A$.

Complementation. Let $A \subset S$. The complement of A, denoted by $\overline{A}$, is the set containing all elements of S that are not in A. That is,

$$\overline{A} = \{k | k \in S \text{ and } k \notin A\}$$

Example 1.2 Let $S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, $A = \{1, 2, 4, 7\}$, and $B = \{1, 3, 4, 6\}$. Then $\overline{A} = \{3, 5, 6, 8, 9, 10\}$, and $\overline{B} = \{2, 5, 7, 8, 9, 10\}$. ▲

Union. The union of two sets A and B, denoted by $A \cup B$, is the set containing all the elements of either A or B or both A and B. That is,

$$A \cup B = \{k | k \in A \text{ or } k \in B\}$$

In Example 1.2, $A \cup B = \{1, 2, 3, 4, 6, 7\}$.

Intersection. The intersection of two sets A and B, denoted by $A \cap B$, is the set containing all the elements that are in both A and B. That is,

$$A \cap B = \{k | k \in A \text{ and } k \in B\}$$

In Example 1.2, $A \cap B = \{1, 4\}$.

Difference. The difference of two sets A and B, denoted by $A - B$, is the set containing all elements of A that are not in B. That is,

$$A - B = \{k | k \in A \text{ and } k \notin B\}$$

Note that $A - B \neq B - A$. From Example 1.2 we find that $A - B = \{2, 7\}$, while $B - A = \{3, 6\}$.

Disjoint Sets. Two sets A and B are called disjoint (or mutually exclusive) sets if they contain no elements in common, which means that $A \cap B = \emptyset$.

1.5.2 Number of Subsets of a Set

Let a set A contain n elements labeled $a_1, a_2, \ldots, a_n$. The number of possible subsets of A is 2^n, which can be obtained as follows for the case of $n = 3$. The eight subsets are given by $\{\overline{a_1}, \overline{a_2}, \overline{a_3}\} = \emptyset$, $\{\overline{a_1}, \overline{a_2}, a_3\}$, $\{\overline{a_1}, a_2, \overline{a_3}\}$, $\{\overline{a_1}, a_2, a_3\}$, $\{a_1, \overline{a_2}, \overline{a_3}\}$, $\{a_1, \overline{a_2}, a_3\}$, $\{a_1, a_2, \overline{a_3}\}$, $\{a_1, a_2, a_3\} = A$; where $\overline{a_k}$ indicates that the element a_k is not included. By convention, if a_k is not an element of a subset, its complement is not explicitly included in the subset. Thus, the subsets are $\emptyset$, $\{a_1\}$, $\{a_2\}$, $\{a_3\}$, $\{a_1, a_2\}$, $\{a_1, a_3\}$, $\{a_2, a_3\}$, $\{a_1, a_2, a_3\} = A$. Since the number of subsets includes the null set, the number of subsets that contain at least one element is $2^n - 1$. The result can be extended to the case of $n > 3$.

The set of all subsets of a set A is called the *power set* of A and denoted by $s(A)$. Thus, for the set $A = \{a, b, c\}$, the power set of A is given by

$$s(A) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$$

The number of members of a set A is called the cardinality of A and denoted by $|A|$. Thus, if the cardinality of the set A is n, then the cardinality of the power set of A is $|s(A)| = 2^n$.

1.5.3 Venn Diagram

The different set operations discussed in the previous section can be graphically represented by the Venn diagram. Figure 1.3 illustrates the complementation, union, intersection, and difference operations on two sets A and B. The universal set is represented by the set of points inside a rectangle. The sets A and B are represented by the sets of points inside circles.

1.5.4 Set Identities

The operations of forming unions, intersections, and complements of sets obey certain rules similar to the rules of algebra. These rules include the following:

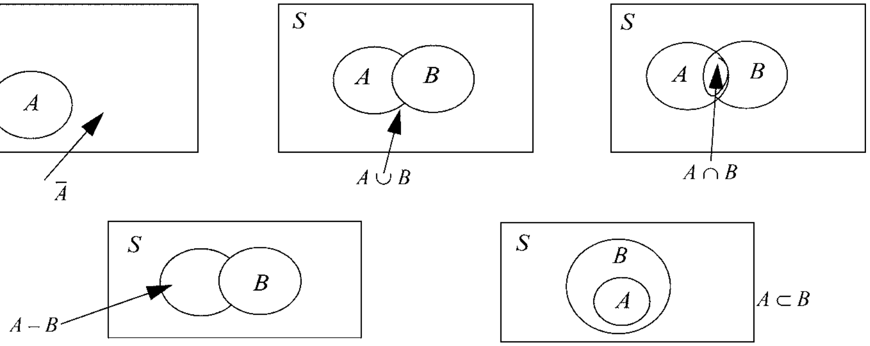

Figure 1.3 Venn Diagrams of Different Set Operations

- *Commutative law for unions*: $A \cup B = B \cup A$, which states that the order of the union operation on two sets is immaterial.
- *Commutative law for intersections:* $A \cap B = B \cap A$, which states that the order of the intersection operation on two sets is immaterial.
- *Associative law for unions*: $A \cup (B \cup C) = (A \cup B) \cup C$, which states that in performing the union operation on three sets, we can proceed in two ways: We can first perform the union operation on the first two sets to obtain an intermediate result and then perform the operation on the result and the third set. The same result is obtained if we first perform the operation on the last two sets and then perform the operation on the first set and the result obtained from the operation on the last two sets.
- *Associative law for intersections*: $A \cap (B \cap C) = (A \cap B) \cap C$, which states that in performing the intersection operation on three sets, we can proceed in two ways: We can first perform the intersection operation on the first two sets to obtain an intermediate result and then perform the operation on the result and the third set. The same result is obtained if we first perform the operation on the last two sets and then perform the operation on the first set and the result obtained from the operation on the last two sets.
- *First distributive law*: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, which states that the intersection of a set A and the union of two sets B and C is equal to the union of the intersection of A and B and the intersection of A and C. This law can be extended as follows:

$$A \cap \left(\bigcup_{i=1}^{n} B_i \right) = \bigcup_{i=1}^{n} (A \cap B_i)$$

- *Second distributive law*: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$, which states that the union of a set A and the intersection of two sets B and C is equal to the intersection of the union of A and B and the union of A and C. The law can also be extended as follows:

$$A \cup \left(\bigcap_{i=1}^{n} B_i \right) = \bigcap_{i=1}^{n} (A \cup B_i)$$

- *De Morgan's first law*: $\overline{A \cup B} = \overline{A} \cap \overline{B}$, which states that the complement of the union of two sets is equal to the intersection of the complements of the sets. The law can be extended to include more than two sets as follows:

$$\overline{\bigcup_{i=1}^{n} A_i} = \bigcap_{i=1}^{n} \overline{A_i}$$

- *De Morgan's second law*: $\overline{A \cap B} = \overline{A} \cup \overline{B}$, which states that the complement of the intersection of two sets is equal to the union of the complements of the sets. The law can also be extended to include more than two sets as follows:

$$\overline{\bigcap_{i=1}^{n} A_i} = \bigcup_{i=1}^{n} \overline{A_i}$$

- Other identities include the following:
  - $A - B = A \cap \overline{B}$, which states that the difference of A and B is equal to the intersection of A and the complement of B.
  - $A \cup S = S$, which states that the union of A and the universal set S is equal to S.
  - $A \cap S = A$, which states that the intersection of A and the universal set S is equal to A.
  - $A \cup \varnothing = A$, which states that the union of A and the null set is equal to A.
  - $A \cap \varnothing = \varnothing$, which states that the intersection of A and the null set is equal to the null set.
  - $\overline{S} = \varnothing$, which states that the complement of the universal set is equal to the null set.
  - For any two sets A and B, $A = (A \cap B) \cup (A \cap \overline{B})$, which states that the set A is equal to the union of the intersection of A and B and the intersection of A and the complement of B.

The way to prove these identities is to show that any point contained in the event on the left side of the equality is also contained in the event on the right side and vice versa.

1.5.5 Duality Principle

The duality principle states that any true result involving sets is also true when we replace unions by intersections, intersections by unions, and sets by their complements, and if we reverse the inclusion symbols $\subset$ and $\supset$. For example, if we replace the union in the first distributive law with intersection and intersection with union, we obtain the second distributive law and vice versa. The same result holds for the two De Morgan's laws.

1.6 Properties of Probability

We now combine the results of set identities with those of the axiomatic definition of probability. (See Section 1.3.1.) From these two sections we obtain the following results:

1. $P(\overline{A}) = 1 - P(A)$, which states that the probability of the complement of A is one minus the probability of A.
2. $P(\emptyset) = 0$, which states that the impossible (or null) event has probability zero.
3. If $A \subset B$, then $P(A) \leq P(B)$. That is, if A is a subset of B, the probability of A is at most the probability of B (or the probability of A cannot exceed the probability of B).
4. $P(A) \leq 1$, which means that the probability of event A is at most 1.
5. If $A = A_1 \cup A_2 \cup \cdots \cup A_n$, where $A_1, A_2, \ldots, A_n$ are mutually exclusive events, then
$$P(A) = P(A_1) + P(A_2) + \cdots + P(A_n)$$
6. For any two events A and B, $P(A) = P(A \cap B) + P(A \cap \overline{B})$, which follows from the set identity: $A = (A \cap B) \cup (A \cap \overline{B})$. Since $A \cap B$ and $A \cap \overline{B}$ are mutually exclusive events, the result follows.
7. For any two events A and B, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. This result can be proved by making use of the Venn diagram. Figure 1.4a represents a Venn diagram in which the left circle represents event A and the right circle represents event B. In Figure 1.4b we divide the diagram into three mutually

Figure 1.4 Venn Diagram of $A \cup B$

exclusive sections labeled I, II, and III, where section I represents all points in A that are not in B, section II represents all points in both A and B, and section III represents all points in B that are not in A.

From Figure 1.4b, we observe that

$$A \cup B = \text{I} \cup \text{II} \cup \text{III}$$
$$A = \text{I} \cup \text{II}$$
$$B = \text{II} \cup \text{III}$$

Since I, II, and III are mutually exclusive, Property 5 implies that

$$P(A \cup B) = P(\text{I}) + P(\text{II}) + P(\text{III})$$
$$P(A) = P(\text{I}) + P(\text{II})$$
$$P(B) = P(\text{II}) + P(\text{III})$$

Thus,

$$\begin{aligned} P(A) + P(B) &= P(\text{I}) + 2P(\text{II}) + P(\text{III}) = \{P(\text{I}) + P(\text{II}) + P(\text{III})\} + P(\text{II}) \\ &= P(A \cup B) + P(\text{II}) \end{aligned}$$

which shows that

$$P(A \cup B) = P(A) + P(B) - P(\text{II}) = P(A) + P(B) - P(A \cap B)$$

8. We can extend Property 7 to the case of three events. If A_1, A_2, A_3 are three events in S, then

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) \\ - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3) \end{aligned}$$

This can be further generalized to the case of n arbitrary events in S as follows:

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \le i \le j \le n} P(A_i \cap A_j) + \sum_{1 \le i \le j \le k \le n} P(A_i \cap A_j \cap A_k) - \cdots$$

That is, to find the probability that at least one of the n events A_i occurs, first add the probability of each event, then subtract the probabilities of all possible two-way intersections, then add the probabilities of all possible three-way intersections, and so on.

1.7 Conditional Probability

Consider the following experiment. We are interested in the sum of the numbers that appear when two dice are tossed. Suppose we are interested in the event that the sum of the two tosses is 7, and we observe that the first toss is 4. Based on this fact, the six possible and equally likely outcomes of the two tosses are {4, 1}, {4, 2}, {4, 3}, {4, 4}, {4, 5}, and {4, 6}. In the absence of the information that the first toss is 4, there would have been 36 sample points in the sample space. But with the information on the outcome of the first toss, there are now only 6 sample points.

Let A denote the event that the sum of the two dice is 7, and let B denote the event that the first die is 4. The conditional probability of event A given event B, denoted by $P(A|B)$, is defined by

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{P(\{4,3\})}{P(\{4,1\}) + P(\{4,2\}) + P(\{4,3\}) + P(\{4,4\}) + P(\{4,5\}) + P(\{4,6\})} \\ &= \frac{(1/36)}{(1/6)} = \frac{1}{6} \end{aligned}$$

Note that $P(A|B)$ is only defined when $P(B) > 0$.

Example 1.3 A bag contains eight red balls, four green balls, and eight yellow balls. A ball is drawn at random from the bag, and it is not a red ball. What is the probability that it is a green ball?

Solution Let G denote the event that the selected ball is a green ball, and let $\overline{R}$ denote the event that it is not a red ball. Then, $P(G) = 4/20 = 1/5$, since there are 4 green balls out of a total of 20 balls, and $P(\overline{R}) = 12/20 = 3/5$, since there are 12 balls out of 20 that are not red. Now,

$$P(G|\overline{R}) = \frac{P(G \cap \overline{R})}{P(\overline{R})}$$

But if the ball is green and not red, it must be green. Thus, we obtain that $G \cap \overline{R} = G$ and

$$P(G|\overline{R}) = \frac{P(G \cap \overline{R})}{P(\overline{R})} = \frac{P(G)}{P(\overline{R})} = \frac{1/5}{3/5} = \frac{1}{3}$$

▲

Example 1.4 A fair coin was tossed two times. Given that the first toss resulted in heads, what is the probability that both tosses resulted in heads?

Solution Because the coin is fair, the four sample points of the sample space $S = \{HH, HT, TH, TT\}$ are equally likely. Let X denote the event that both tosses came up heads; that is, $X = \{HH\}$. Let Y denote the event that the first toss came up heads; that is, $Y = \{HH, HT\}$. The probability that both tosses resulted in heads, given that the first toss resulted in heads, is given by

$$P(X|Y) = \frac{P(X \cap Y)}{P(Y)} = \frac{P(X)}{P(Y)} = \frac{1/4}{2/4} = \frac{1}{2}$$

▲

1.7.1 Total Probability and the Bayes' Theorem

A partition of a set A is a set $\{A_1, A_2, \ldots, A_n\}$ with the following properties:

a. $A_i \subseteq A$, $i = 1, 2, \ldots, n$, which means that A is a set of subsets.

b. $A_i \cap A_k = \emptyset$, $i = 1, 2, \ldots, n$; $k = 1, 2, \ldots, n$; $i \neq k$, which means that the subsets are mutually (or pairwise) disjoint; that is, no two subsets have any element in common.

c. $A_1 \cup A_2 \cup \cdots \cup A_n = A$, which means that the subsets are collectively exhaustive. That is, the subsets together include all possible values of the set A.

Proposition 1.1. *Let* $\{A_1, A_2, \ldots, A_n\}$ *be a partition of the sample space S, and suppose each one of the events* $A_1, A_2, \ldots, A_n$, *has nonzero probability of occurrence. Let* A *be any event. Then*

$$\begin{aligned} P(A) &= P(A_1)P(A|A_1) + P(A_2)P(A|A_2) + \cdots + P(A_n)P(A|A_n) \\ &= \sum_{i=1}^{n} P(A_i)P(A|A_i) \end{aligned}$$

Proof. The proof is based on the observation that because $\{A_1, A_2, \ldots, A_n\}$ is a partition of S, the set $\{A \cap A_1, A \cap A_2, \ldots, A \cap A_n\}$ is a partition of the event A because if A occurs, then it must occur in conjunction with one of the A_i's. Thus, we can express A as the union of n mutually exclusive events. That is,

$$A = (A \cap A_1) \cup (A \cap A_2) \cup \cdots \cup (A \cap A_n)$$

Since these events are mutually exclusive, we obtain

$$P(A) = P(A \cap A_1) + P(A \cap A_2) + \cdots + P(A \cap A_n)$$

From our definition of conditional probability, $P(A \cap A_i) = P(A_i)P(A|A_i)$, which exists because we assumed in the proposition that the events $A_1, A_2, \ldots, A_n$ have nonzero probabilities. Substituting the definition of conditional probabilities, we obtain the desired result:

$$P(A) = P(A_1)P(A|A_1) + P(A_2)P(A|A_2) + \cdots + P(A_n)P(A|A_n)$$

The preceding result is defined as the *total probability* of event A, which will be useful in the remainder of the book. ■

Example 1.5 A student buys 1000 integrated circuits (ICs) from supplier A, 2000 ICs from supplier B, and 3000 ICs from supplier C. He tested the ICs and found that the conditional probability of an IC being defective depends on the supplier from whom it was bought. Specifically, given that an IC came from supplier A, the probability that it is defective is 0.05; given that an IC came from supplier B, the probability that it is defective is 0.10; and given that an IC came from supplier C, the probability that it is defective is 0.10. If the ICs from the three suppliers are mixed together and one is selected at random, what is the probability that it is defective?

Solution Let $P(A)$, $P(B)$, and $P(C)$ denote the probability that a randomly selected IC came from supplier A, B, and C, respectively. Also, let $P(D|A)$ denote the conditional probability that an IC is defective, given that it came from supplier A; $P(D|B)$ denote the conditional probability that an IC is defective, given that it came from supplier B; and $P(D|C)$ denote the conditional probability that an IC is defective, given that it came from supplier C. Then the following are true:

$$\begin{aligned}
P(D|A) &= 0.05 \\
P(D|B) &= 0.10 \\
P(D|C) &= 0.10 \\
P(A) &= \frac{1000}{1000+2000+3000} = \frac{1}{6} \\
P(B) &= \frac{2000}{1000+2000+3000} = \frac{1}{3} \\
P(C) &= \frac{3000}{1000+2000+3000} = \frac{1}{2}
\end{aligned}$$

Let $P(D)$ denote the unconditional probability that a randomly selected IC is defective. Then, from the principles of total probability,

$$\begin{aligned}
P(D) &= P(D|A)P(A) + P(D|B)P(B) + P(D|C)P(C) \\
&= (0.05)(1/6) + (0.10)(1/3) + (0.10)(1/2) \\
&= 0.09167
\end{aligned}$$

▲

We now go back to the general discussion. Suppose event A has occurred, but we do not know which of the mutually exclusive and exhaustive events $A_1, A_2, \ldots, A_n$ holds true. The conditional probability that event A_k occurred, given that A occurred, is given by

$$P(A_k|A) = \frac{P(A_k \cap A)}{P(A)} = \frac{P(A_k \cap A)}{\sum_{i=1}^{n} P(A|A_i)P(A_i)}$$

where the second equality follows from the total probability of event A. Since $P(A_k \cap A) = P(A|A_k)P(A_k)$, the preceding equation can be rewritten as follows:

$$P(A_k|A) = \frac{P(A_k \cap A)}{P(A)} = \frac{P(A|A_k)P(A_k)}{\sum_{i=1}^{n} P(A|A_i)P(A_i)}$$

This result is called the *Bayes' formula* (or *Bayes' rule*).

Example 1.6 In Example 1.5, given that a randomly selected IC is defective, what is the probability that it came from supplier A?

Solution Using the same notation as in Example 1.5, the probability that the randomly selected IC came from supplier A, given that it is defective, is given by

$$\begin{aligned} P(A|D) &= \frac{P(D \cap A)}{P(D)} \\ &= \frac{P(D|A)P(A)}{P(D|A)P(A) + P(D|B)P(B) + P(D|C)P(C)} \\ &= \frac{(0.05)(1/6)}{(0.05)(1/6) + (0.10)(1/3) + (0.10)(1/2)} \\ &= 0.0909 \end{aligned}$$

▲

Example 1.7 (***The Binary Symmetric Channel***) A discrete channel is characterized by an input alphabet $X = \{x_1, x_2, \ldots, x_n\}$; an output alphabet $Y = \{y_1, y_2, \ldots, y_m\}$; and a set of conditional probabilities (called *transition probabilities*), P_{ij}, which are defined as follows: $P_{ij} = P(y_j|x_i) = P[\text{receiving symbol } y_j | \text{symbol } x_i \text{ was transmitted}]$, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, m$. The binary channel is a special case of the discrete channel, where $n = m = 2$. It can be represented as shown in Figure 1.5.

In the binary channel, an error occurs if y_2 is received when x_1 is transmitted or y_1 is received when x_2 is transmitted. Thus, the probability of error, P_e, is given by

$$\begin{aligned} P_e &= P(x_1 \cap y_2) + P(x_2 \cap y_1) \\ &= P(x_1)P(y_2|x_1) + P(x_2)P(y_1|x_2) \\ &= P(x_2)P_{12} + P(x_2)P_{21} \end{aligned}$$

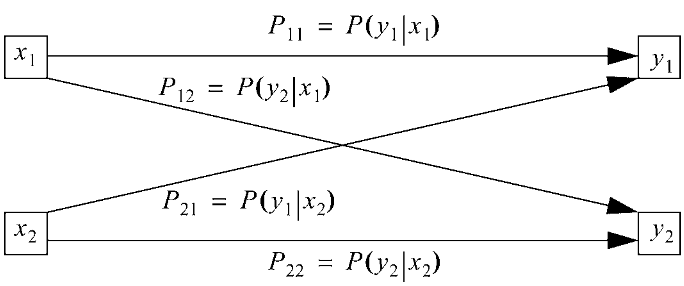

Figure 1.5 The Binary Channel

If $P_{12} = P_{21}$, we say that the channel is a *binary symmetrical channel* (BSC). Also, if in the BSC $P(x_1) = p$, then $P(x_2) = 1 - p = q$.

Consider the BSC shown in Figure 1.6, with $P(x_1) = 0.6$ and $P(x_2) = 0.4$. Evaluate the following:

a. The probability that x_1 was transmitted, given that y_2 was received
b. The probability that x_2 was transmitted, given that y_1 was received
c. The probability that x_1 was transmitted, given that y_1 was received
d. The probability that x_2 was transmitted, given that y_2 was received
e. The unconditional probability of error

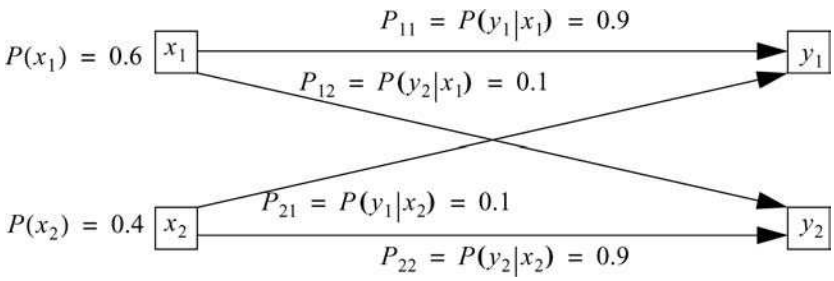

Figure 1.6 The Binary Symmetric Channel for Example 1.7

Solution Let $P(y_1)$ denote the probability that y_1 was received and $P(y_2)$ the probability that y_2 was received. Then

(a) The probability that x_1 was transmitted, given that y_2 was received, is given by

$$
\begin{aligned}
P(x_1|y_2) &= \frac{P(x_1 \cap y_2)}{P(y_2)} = \frac{P(y_2|x_1)P(x_1)}{P(y_2|x_1)P(x_1) + P(y_2|x_2)P(x_2)} \\
&= \frac{(0.1)(0.6)}{(0.1)(0.6) + (0.9)(0.4)} \\
&= 0.143
\end{aligned}
$$

(b) The probability that x_2 was transmitted, given that y_1 was received, is given by

$$
P(x_2|y_1) = \frac{P(x_2 \cap y_1)}{P(y_1)} = \frac{P(y_1|x_2)P(x_2)}{P(y_1|x_1)P(x_1) + P(y_1|x_2)P(x_2)}
$$

$$= \frac{(0.1)(0.4)}{(0.9)(0.6) + (0.1)(0.4)}$$
$$= 0.069$$

(c) The probability that x_1 was transmitted, given that y_1 was received, is given by

$$P(x_1|y_1) = \frac{P(x_1 \cap y_1)}{P(y_1)} = \frac{P(y_1|x_1)P(x_1)}{P(y_1|x_1)P(x_1) + P(y_1|x_2)P(x_2)}$$
$$= \frac{(0.9)(0.6)}{(0.9)(0.6) + (0.1)(0.4)}$$
$$= 0.931$$

(d) The probability that x_2 was transmitted, given that y_2 was received, is given by

$$P(x_2|y_2) = \frac{P(x_2 \cap y_2)}{P(y_2)} = \frac{P(y_2|x_2)P(x_2)}{P(y_2|x_1)P(x_1) + P(y_2|x_2)P(x_2)}$$
$$= \frac{(0.9)(0.4)}{(0.1)(0.6) + (0.9)(0.4)}$$
$$= 0.857$$

(e) The unconditional probability of error is given by

$$P_e = P(x_1)P_{12} + P(x_2)P_{21}$$
$$= (0.6)(0.1) + (0.4)(0.1)$$
$$= 0.1$$

▲

Example 1.8 The quarterback for a certain football team has a good game with probability 0.6 and a bad game with probability 0.4. When he has a good game, he throws at least one interception with a probability of 0.2; and when he has a bad game, he throws at least one interception with a probability of 0.5. Given that he threw at least one interception in a particular game, what is the probability that he had a good game?

Solution Let G denote the event that the quarterback has a good game and B the event that he had a bad game. Similarly, let I denote the event that he throws at least one interception. Then we have that

$$\begin{aligned}
P(G) &= 0.6 \\
P(B) &= 0.4 \\
P(I|G) &= 0.2 \\
P(I|B) &= 0.5 \\
P(G|I) &= \frac{P(G \cap I)}{P(I)}
\end{aligned}$$

According to the Bayes' formula, the last equation becomes

$$\begin{aligned}
P(G|I) &= \frac{P(G \cap I)}{P(I)} = \frac{P(I|G)P(G)}{P(I|G)P(G) + P(I|B)P(B)} \\
&= \frac{(0.2)(0.6)}{(0.2)(0.6) + (0.5)(0.4)} = \frac{0.12}{0.32} \\
&= 3/8 = 0.375
\end{aligned}$$

▲

Example 1.9 Two events A and B are such that $P[A \cap B] = 0.15$, $P[A \cup B] = 0.65$, and $P[A|B] = 0.5$. Find $P[B|A]$.

Solution $P[A \cup B] = P[A] + P[B] - P[A \cap B] \Rightarrow 0.65 = P[A] + P[B] - 0.15$. This means that $P[A] + P[B] = 0.65 + 0.15 = 0.80$. Also, $P[A \cap B] = P[B] \times P[A|B]$. This then means that

$$P[B] = \frac{P[A \cap B]}{P[A|B]} = \frac{0.15}{0.50} = 0.30$$

Thus, $P[A] = 0.80 - 0.30 = 0.50$. Since $P[A \cap B] = P[A] \times P[B|A]$, we have that

$$P[B|A] = \frac{P[A \cap B]}{P[A]} = \frac{0.15}{0.50} = 0.30$$

▲

Example 1.10 A student went to the post office to mail a package to his parents. He gave the postal attendant a bill he believed was \$20. However, the postal attendant gave him change based on her belief that she received a \$10 bill from the student. The student started to dispute the change. Both the student and the postal attendant are honest but may make mistakes. If the postal attendant's drawer contains 30 \$20 bills and 20 \$10 bills, and she correctly identifies bills 90% of the time, what is the probability that the student's claim is valid?

Solution Let A denote the event that the student gave the postal attendant a \$10 bill and B the event that the student gave the postal attendant a \$20 bill. Let V denote the event that the student's claim is valid. Finally, let L denote the event that the postal attendant said that the student gave her a \$10 bill. Since there are 30 \$20 bills and 20 \$10 bills in the drawer, the probability that the money the student gave the postal attendat was a \$20 bill is $30/(20+30)=0.6$, and the probability that it was a \$10 bill is $1-0.6=0.4$. Thus,

$$\begin{aligned} P(L) &= P(L|A)P(A) + P(L|B)P(B) \\ &= 0.90 \times 0.4 + 0.10 \times 0.6 = 0.42 \end{aligned}$$

Thus, the probability that the student's claim is valid is the probability that he gave the postal attendant a \$20 bill, given that she said that he gave her a \$10 bill. Using Bayes' formula we obtain

$$\begin{aligned} P(V|L) &= \frac{P(V\cap L)}{P(L)} = \frac{P(L|V)P(V)}{P(L)} \\ &= \frac{0.10 \times 0.60}{0.42} = \frac{1}{7} = 0.1428 \end{aligned}$$

▲

Example 1.11 An aircraft maintenance company bought equipment for detecting structural defects in aircrafts. Tests indicate that 95% of the time the equipment detects defects when they actually exist, and 1% of the time it gives a false alarm that indicates the presence of a structural defect when in fact there is none. If 2% of the aircrafts actually have structural defects, what is the probability that an aircraft actually has a structural defect given that the equipment indicates that it has a structural defect?

Solution Let D denote the event that an aircraft has a structural defect and B the event that the test indicates that there is a structural defect. Then we are required to find $P(D|B)$. Using Bayes' formula we obtain

$$\begin{aligned} P(D|B) &= \frac{P(D\cap B)}{P(B)} = \frac{P(B|D)P(D)}{P(B|D)P(D) + P(B|\overline{D})P(\overline{D})} \\ &= \frac{0.95 \times 0.02}{\{0.95 \times 0.02\} + \{0.01 \times 0.98\}} = 0.660 \end{aligned}$$

Thus, only 66% of the aircrafts that the equipment diagnoses as having structural defects actually have structural defects. ▲

1.7.2 Tree Diagram

Conditional probabilities are used to model experiments that take place in stages. The outcomes of such experiments are conveniently represented by a tree diagram. A tree is a connected graph that contains no circuit (or loop). Every two nodes in the tree have a unique path connecting them. Line segments called branches interconnect the nodes. Each branch may split into other branches, or it may terminate. When used to model an experiment, the nodes of the tree represent events of the experiment. The number of branches that emanate from a node represents the number of events that can occur, given that the event represented by that node occurs. The node that has no predecessor is called the *root* of the tree, and any node that has no successor or children is called a *leaf* of the tree. The events of interest are usually defined at the leaves by tracing the outcomes of the experiment from the root to each leaf.

The conditional probabilities appear on the branches leading from the node representing an event to the nodes representing the next events of the experiment. A path through the tree corresponds to a possible outcome of the experiment. Thus, the product of all the branch probabilities from the root of the tree to any node is equal to the probability of the event represented by that node.

Consider an experiment that consists of three tosses of a coin. Let p denote the probability of heads in a toss; then $1-p$ is the probability of tails in a toss. Figure 1.7 is the tree diagram for the experiment.

Let A be the event "the first toss came up heads," and let B be the event "the second toss came up tails." Then from Figure 1.7, $P(A)=p$ and $P(B)=1-p$. Since $P(A\cap B)=p(1-p)$, we have that

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

We can obtain the same result by noting that the event $A\cup B$ consists of the following six-element set:

$$A\cup B=\{HHH, HHT, HTH, HTT, TTH, TTT\}$$

Example 1.12 A university has twice as many undergraduate students as graduate students. Twenty five percent of the graduate students live on campus, and 10% of the undergraduate students live on campus.

a. If a student is chosen at random from the student population, what is the probability that the student is an undergraduate student living on campus?

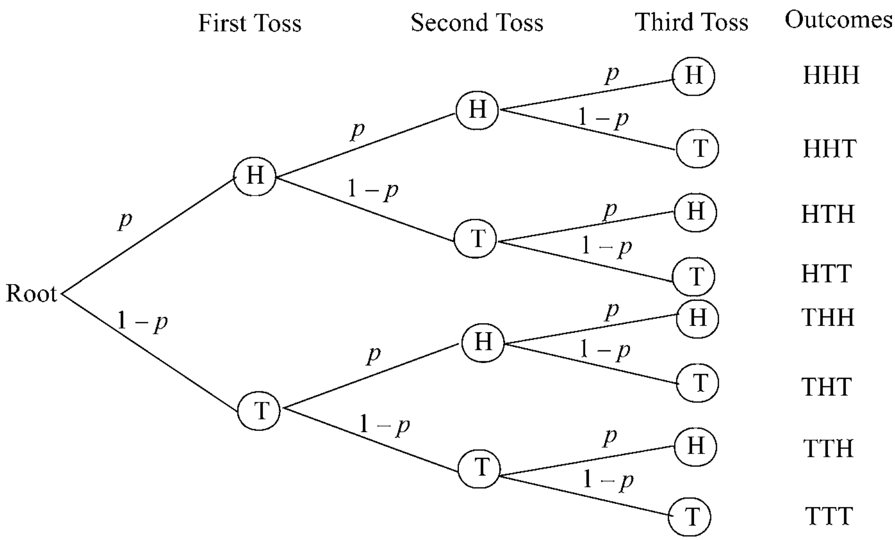

Figure 1.7 Tree Diagram for Three Tosses of a Coin

b. If a student living on campus is chosen at random, what is the probability that the student is a graduate student?

Solution We use the tree diagram to solve the problem. Since there are twice as many undergraduate students as there are graduate students, the proportion of undergraduate students in the population is 2/3, and the proportion of graduate students is 1/3. These as well as the other data are shown as the labels on the branches of the tree in Figure 1.8. In the figure G denotes graduate students, U denotes undergraduate students, ON denotes living on campus, and OFF denotes living off campus.

(a) From the figure we see that the probability that a randomly selected student is an undergraduate student living on campus is 0.067. We can also solve the problem directly as follows. We are required to find the probability of choosing an undergraduate student who lives on campus, which is $P(U \cap ON)$. This is given by

$$P(U \cap ON) = P(ON|U)P(U) = 0.10 \times \frac{2}{3} = 0.067$$

(b) From the tree, the probability that a student lives on campus is $(0.067 + 0.083)$. Thus, the probability that a randomly selected student living on campus is a graduate student is $0.083/(0.083 + 0.067) = 0.55$. We can also use the

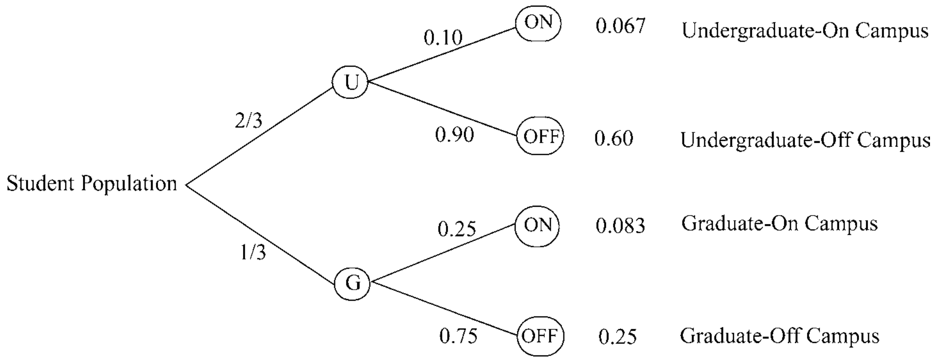

Figure 1.8 Figure for Example 1.12

Bayes' theorem to solve the problem as follows:

$$P(G|ON) = \frac{P(G \cap ON)}{P(ON)} = \frac{P(ON|G)P(G)}{P(ON|G)P(G) + P(ON|U)P(U)}$$
$$= \frac{(0.25)(1/3)}{(0.25)(1/3) + (0.1)(2/3)} = \frac{5}{9}$$
$$= 0.55$$

▲

1.8 Independent Events

Two events A and B are defined to be independent if the knowledge that one has occurred does not change or affect the probability that the other will occur. In particular, if events A and B are independent, the conditional probability of event A, given event B, $P(A|B)$, is equal to the probability of event A. That is, events A and B are independent if

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

Since by definition $P(A \cap B) = P(A|B)P(B)$, an alternative definition of the independence of events is that events A and B are independent if

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

The definition of independence can be extended to multiple events. The n events $A_1, A_2, \ldots, A_n$ are said to be independent if the following conditions are true:

$$P(A_i \cap A_j) = P(A_i)P(A_j)$$

$$P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k)$$
$$\cdots$$
$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2)\ldots P(A_n)$$

This is true for all $1 \leq i < j < k < \cdots \leq n$. That is, these events are pairwise independent, independent in triplets, and so on.

Example 1.13 A red die and a blue die are rolled together. What is the probability that we obtain 4 on the red die and 2 on the blue die?

Solution Let R denote the event "4 on the red die," and let B denote the event "2 on the blue die." We are, therefore, required to find $P(R \cap B)$. Since the outcome of one die does not affect the outcome of the other die, the events R and B are independent. Thus, since $P(R) = 1/6$ and $P(B) = 1/6$, $P(R \cap B) = P(R)P(B) = 1/36$. ▲

Example 1.14 Two coins are tossed. Let A denote the event "at most one head on the two tosses," and let B denote the event "one head and one tail in both tosses." Are A and B independent events?

Solution The sample space of the experiment is $S = \{HH, HT, TH, TT\}$. Now, events are defined as follows: $A = \{HT, TH, TT\}$ and $B = \{HT, TH\}$. Also, $A \cap B = \{HT, TH\}$. Thus,

$$P(A) = \frac{3}{4}$$
$$P(B) = \frac{2}{4} = \frac{1}{2}$$
$$P(A \cap B) = \frac{2}{4} = \frac{1}{2}$$
$$P(A)P(B) = \frac{8}{3}$$

Since $P(A \cap B) \neq P(A)P(B)$, we conclude that events A and B are not independent. ▲

Proposition 1.2. *If* A *and* B *are independent events, then so are events* A *and* $\overline{\mathrm{B}}$, *events* $\overline{\mathrm{A}}$ *and* B, *and events* $\overline{\mathrm{A}}$ *and* $\overline{\mathrm{B}}$.

Proof. Event A can be written as follows: $A = (A \cap B) \cup (A \cap \overline{B})$. Since the events $A \cap B$ and $A \cap \overline{B}$ are mutually exclusive, we may write

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

where the last equality follows from the fact that A and B are independent. Thus, we obtain

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

which proves that events A and $\overline{B}$ are independent. To prove that events $\overline{A}$ and B are independent, we start with $B = (A \cap B) \cup (\overline{A} \cap B)$. Using the same fact that the two events are mutually exclusive, we derive the condition for independence. Finally, to prove that events $\overline{A}$ and $\overline{B}$ are independent, we start with $\overline{A} = (\overline{A} \cap B) \cup (\overline{A} \cap \overline{B})$ and proceed as previously using the results already established. ■

Example 1.15 A and B are two independent events defined in the same sample space. They have the following probabilities: $P[A] = x$ and $P[B] = y$. Find the probabilities of the following events in terms of x and y:

a. Neither event A nor event B occurs
b. Event A occurs but event B does not occur
c. Either event A occurs or event B does not occur

Solution Since events A and B are independent, we know from Proposition 1.2 that events A and $\overline{B}$ are independent, events $\overline{A}$ and B are independent, and events $\overline{A}$ and $\overline{B}$ are also independent.

(a) The probability that neither event A nor event B occurs is the probability that event A does not occur and event B does not occur, which is given by

$$P_{\overline{a}\overline{b}} = P(\overline{A} \cap \overline{B}) = P(\overline{A})P(\overline{B}) = (1 - x)(1 - y)$$

where the second equality is due to independence of $\overline{A}$ and $\overline{B}$.

(b) The probability that event A occurs but event B does not occur is the probability that event A occurs and event B does not occur, which is given by

$$P_{a\overline{b}} = P(A \cap \overline{B}) = P(A)P(\overline{B}) = x(1 - y)$$

where the second equality is due to the independence of A and $\overline{B}$.

(c) The probability that either event A occurs or event B does not occur is given by

$$\begin{aligned} P(A \cup \overline{B}) &= P(A) + P(\overline{B}) - P(A \cap \overline{B}) \\ &= P(A) + P(\overline{B}) - P(A)P(\overline{B}) = x + (1-y) - x(1-y) \\ &= 1 - y(1-x) \end{aligned}$$

where the second equality is due to the independence of A and $\overline{B}$.

▲

Example 1.16 Jim and Bill like to shoot at targets. Jim can hit a target with a probability of 0.8, while Bill can hit a target with a probability of 0.7. If both fire at a target at the same time, what is the probability that the target is hit at least once?

Solution Let J denote the event that Jim hits a target and B the event that Bill hits a target. Since the outcome of Bill's shot is not affected by the outcome of Jim's shot, and vice versa, the events J and B are independent. Because B and J are independent events, the events J and $\overline{B}$ are independent, and the events B and $\overline{J}$ are independent. Thus, the probability that the target is hit at least once is the probability of the union of its being hit once and its being hit twice. That is, if p is the probability that the target is hit at least once, then

$$\begin{aligned} p &= P(\{J \cap \overline{B}\} \cup \{\overline{J} \cap B\} \cup \{J \cap B\}) = P(J \cap \overline{B}) + P(\overline{J} \cap B) + P(J \cap B) \\ &= P(J)P(\overline{B}) + P(\overline{J})P(B) + P(J)P(B) \\ &= (0.8)(0.3) + (0.2)(0.7) + (0.8)(0.7) \\ &= 0.94 \end{aligned}$$

▲

1.9 Combined Experiments

Until now our discussion has been limited to single experiments. Sometimes we are required to form an experiment by combining multiple individual experiments. Consider the case of two experiments in which one experiment has the sample space S_1 with N sample points, and the other has the sample space S_2 with M sample points. That is,

$$\begin{aligned} S_1 &= \{x_1, x_2, \ldots, x_N\} \\ S_2 &= \{y_1, y_2, \ldots, y_M\} \end{aligned}$$

If we form an experiment that is a combination of these two experiments, the sample space of the combined experiment is called the *combined space* (or the *Cartesian product space*) and is defined by

$$S = S_1 \times S_2 = \{(x_i, y_j) | x_i \in S_1, y_j \in S_2, i = 1, 2, \ldots, N; j = 1, 2, \ldots, M\}$$

The combined sample space of an experiment that is a combination of N experiments with sample spaces S_k, $k = 1, 2, \ldots, N$, is given by

$$S = S_1 \times S_2 \times \cdots \times S_N$$

Note that if L_k is the number of sample points in S_k, $k = 1, 2, \ldots, N$, then the number of sample points in S (also called the *cardinality* of S) is given by $L = L_1 \times L_2 \times \cdots \times L_N$. That is, the cardinality of S is the product of the cardinalities of the sample spaces of the different experiments.

Example 1.17 Consider a combined experiment formed from two experiments. The first experiment consists of tossing a coin and the second experiment consists of rolling a die. Let S_1 denote the sample space of the first experiment, and let S_2 denote the sample space of the second experiment. If S denotes the sample space of the combined experiment, we obtain the following:

$$\begin{aligned} S_1 &= \{H, T\} \\ S_2 &= \{1, 2, 3, 4, 5, 6\} \\ S &= \{(H, 1), (H, 2), (H, 3), (H, 4), (H, 5), (H, 6), \\ &\quad (T, 1), (T, 2), (T, 3), (T, 4), (T, 5), (T, 6)\} \end{aligned}$$

As we can see, the number of sample points in S is the product of the number of sample points in the two sample spaces. If we assume that the coin and die are fair, then the sample points in S are equiprobable; that is, each sample point is equally likely to occur. Thus, for example, if we define X to be the event "a head on the coin and an even number of the die," then X and its probability are given by

$$\begin{aligned} X &= \{(H, 2), (H, 4), (H, 6)\} \\ P(X) &= \frac{3}{12} = \frac{1}{4} \end{aligned}$$

An alternative way to solve the problem is as follows. Let H denote the event that the coin comes up heads, and E the event that the die comes up an even number. Then $X = H \cap E$. Because the events H and E are independent, we obtain

$$\begin{aligned} P(X) &= P(H \cap E) = P(H)P(E) \\ &= \frac{1}{2} \times \frac{3}{6} = \frac{1}{4} \end{aligned}$$

▲

1.10 Basic Combinatorial Analysis

Combinatorial analysis deals with counting the number of different ways in which an event of interest can occur. Two basic aspects of combinatorial analysis that are used in probability theory are permutation and combination.

1.10.1 Permutations

Sometimes we are interested in how the outcomes of an experiment can be arranged. For example, if the possible outcomes are A, B, and C, we can think of six possible arrangements of these outcomes: ABC, ACB, BAC, BCA, CAB, and CBA. Each of these arrangements is called a *permutation*. Thus, there are six permutations of a set of three distinct objects. This number can be derived as follows: There are three ways of choosing the first object; after the first object has been chosen, there are two ways of choosing the second object; and after the first two objects have been chosen, there is one way to choose the third object. This means that there are $3 \times 2 \times 1 = 6$ permutations.

For a system of n distinct objects we can apply a similar reasoning to obtain the following number of permutations:

$$n \times (n-1) \times (n-2) \times \cdots \times 3 \times 2 \times 1 = n!$$

where $n!$ is read as "n factorial." By convention, $0! = 1$.

Assume that we want to arrange r of the n objects at a time. The problem now becomes that of finding how many possible sequences of r objects we can get from n objects, where $r \leq n$. This number is denoted by $P(n, r)$ and defined as follows:

$$P(n,r) = \frac{n!}{(n-r)!} = n \times (n-1) \times (n-2) \times \cdots \times (n-r+1) \quad r = 1, 2, \ldots, n$$

The number $P(n, r)$ represents the number of permutations (or sequences) of r objects taken from n objects when the arrangement of the objects within a given sequence is important. Note that when $r = n$, we obtain

$$P(n,n) = n \times (n-1) \times (n-2) \times \cdots \times 3 \times 2 \times 1 = n!$$

Example 1.18 A little girl has six building blocks and is required to select four of them at a time to build a model. If the order of the blocks in each model is important, how many models can she build?

Solution Since the order of objects is important, this is a permutation problem. Therefore, the number of models is given by

$$P(6,4) = \frac{6!}{(6-4)!} = \frac{6!}{2!} = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{2 \times 1} = 360$$

▲

Note that if the little girl were to select three blocks at a time, the number of permutations decreases to 120.

Example 1.19 How many words can be formed from the word SAMPLE? Assume that a formed word does not have to be an actual English word, but it may contain at most as many instances of a letter as there are in the original word (for example, "maa" is not acceptable, since "a" does not appear twice in SAMPLE, but "mas" is allowed).

Solution The words can be single-letter words, two-letter words, three-letter words, four-letter words, five-letter words, or six-letter words. Since the letters of the word SAMPLE are all unique, there are $P(6, k)$ ways of forming k-letter words, $k = 1, 2, \ldots, 6$. Thus, the number of words that can be formed is

$$\begin{aligned} N &= P(6,1) + P(6,2) + P(6,3) + P(6,4) + P(6,5) + P(6,6) \\ &= 6 + 30 + 120 + 360 + 720 + 720 = 1956 \end{aligned}$$

▲

We present the following theorem without proof.

Theorem. *Given a population of* n *elements, let* $n_1, n_2, \ldots, n_k$ *be positive integers such that* $n_1 + n_2 + \cdots + n_k = n$. *Then there are*

$$N = \frac{n!}{n_1! \times n_2! \times \cdots \times n_k!}$$

ways to partition the population into k *subgroups of sizes* $n_1, n_2, \ldots, n_k$, *respectively.*

Example 1.20 Five identical red blocks, two identical white blocks, and three identical blue blocks are arranged in a row. How many different arrangements are possible?

Solution In this example, $n = 5 + 2 + 3 = 10$, $n_1 = 5$, $n_2 = 2$, and $n_3 = 3$. Thus, the number of possible arrangements is given by

$$N = \frac{10!}{5! \times 2! \times 3!} = 2520$$

▲

Example 1.21 How many words can be formed by using all the letters of the word MISSISSIPPI?

Solution The word contains 11 letters consisting of 1 M, 4 S's, 4 I's, and 2 P's. Thus, the number of words that can be formed is

$$N = \frac{11!}{1! \times 4! \times 4! \times 2!} = 34650$$

▲

1.10.2 Circular Arrangement

Consider the problem of seating n people in a circle. Assume that the positions are labeled $1, 2, \ldots, n$. Then, if after one arrangement everyone moves one place to the left or right, that will also be a new arrangement because each person is occupying a new location. However, each person's previous neighbors to the left and right are still his/her neighbors in the new arrangement. This means that such a move does not lead to a new valid arrangement. To solve this problem, one person must remain fixed while others move.

Thus, the number of people being arranged is $n - 1$, which means that the number of possible arrangements is $(n - 1)!$ For example, the number of ways that 10 people can be seated in a circle is $(10 - 1)! = 9! = 362880$.

1.10.3 Applications of Permutations in Probability

Consider a system that contains n distinct objects labeled $a_1, a_2, \ldots, a_n$. Assume that we choose r of these objects in the following manner. We choose the first object, record its type, and put it back into the "population." We then choose the second object, record its type, and put it back into the population. We continue this process until we have chosen a total of r objects. This gives an "ordered sample" consisting of r of the n objects. The question is to determine the number of distinct ordered samples that can be obtained, where two ordered samples are said to be distinct if they differ in at least one entry in a particular position within the samples. Since the number of ways of choosing an object in each round is n, the total number of distinct samples is $n \times n \times \cdots \times n = n^r$.

Assume now that the sampling is done without replacement. That is, after an object has been chosen, it is not put back into the population. Then the next object from the remainder of the population is chosen and not replaced, and so on until all the r objects have been chosen. The total number of possible ways of making this sampling can be obtained by noting that there are n ways to choose the first object, $n-1$ ways to choose the second object, $n-2$ ways to choose the third object, and so on, and finally there are $n-r+1$ ways to choose the rth object. Thus, the total number of distinct samples is $n \times (n-1) \times (n-2) \times \cdots \times (n-r+1)$.

Example 1.22 A subway train consists of n cars. The number of passengers waiting to board the train is $k < n$ and each passenger enters a car at random. What is the probability that all the k passengers end up in different cars of the train?

Solution Without any restriction on the occupancy of the cars, each of the k passengers can enter any one of the n cars. Thus, the number of distinct, unrestricted arrangements of the passengers in the cars is $N = n \times n \times \cdots \times n = n^k$.

If the passengers enter the cars in such a way that there is no more than one passenger in a car, then the first passenger can enter any one of the n cars. After the first passenger has entered a car, the second passenger can enter any one of the $n-1$ remaining cars. Similarly, the third passenger can enter any one of the $n-2$ remaining cars, and so on. Finally, the kth passenger can enter any one of the $n-k+1$ remaining cars. Thus, the total number of distinct arrangements of passengers when no two passengers can be in the same car is $M = n \times (n-1) \times (n-2) \times \cdots \times (n-k+1)$. Therefore, the probability of this event is

$$P = \frac{M}{N} = \frac{n \times (n-1) \times (n-2) \times \cdots \times (n-k+1)}{n^k}$$

▲

Example 1.23 Ten books are placed in random order on a bookshelf. Find the probability that three given books are placed side by side.

Solution The number of unrestricted ways of arranging the books is 10!. Consider the three books to be tied together as a "superbook," which means that there are eight books on the bookshelf including the superbook. The number of ways of arranging these books is 8!. In each of these arrangements the three books can be arranged among themselves in $3! = 6$ ways. Thus, the total number of arrangements with the three books together is 8!3!, and the required probability p is given by

$$p = \frac{8!3!}{10!} = \frac{6 \times 8!}{10 \times 9 \times 8!} = \frac{6}{90} = \frac{1}{15}$$

▲

1.10.4 Combinations

In permutations, the order of objects within a selection is important; that is, the arrangement of objects within a selection is very important. Thus, the arrangement ABC is different from the arrangement ACB even though they both contain the same three objects. In some problems, the order of objects within a selection is not relevant. For example, consider a student who is required to select four subjects out of six subjects in order to graduate. Here, the order of subjects is not important; all that matters is that the student selects four subjects.

Since the order of the objects within a selection is not important, the number of ways of choosing r objects from n objects will be smaller than when the order is important. The number of ways of selecting r objects at a time from n objects when the order of the objects is not important is called the *combination* of r objects taken from n distinct objects and denoted by $C(n, r)$. It is defined as follows:

$$C(n, r) = \binom{n}{r} = \frac{P(n, r)}{r!} = \frac{n!}{(n-r)!r!}$$

Recall that $r!$ is the number of permutations of r objects taken r at a time. Thus, $C(n, r)$ is equal to the number of permutations of n objects taken r at a time divided by the number of permutations of r objects taken r at a time.

Observe that $C(n, r) = C(n, n-r)$, as can be seen from the preceding equation. One very useful combinatorial identity is the following:

$$\binom{n+m}{k} = \sum_{i=0}^{k} \binom{n}{i}\binom{m}{k-i}$$

This identity can easily be proved by considering how many ways we can select k people from a group of m boys and n girls. In particular, when $m = k = n$ we have that

$$\begin{aligned}\binom{2n}{n} &= \binom{n}{0}\binom{n}{n} + \binom{n}{1}\binom{n}{n-1} + \binom{n}{2}\binom{n}{n-2} + \cdots \\ &\quad + \binom{n}{k}\binom{n}{n-k} + \cdots + \binom{n}{n}\binom{n}{0} \\ &= \binom{n}{0}^2 + \binom{n}{1}^2 + \binom{n}{2}^2 + \cdots + \binom{n}{k}^2 + \cdots + \binom{n}{n}^2\end{aligned}$$

where the last equality follows from the fact that $\binom{n}{k} = \binom{n}{n-k}$.

Example 1.24 Evaluate $\binom{16}{8}$.

Solution

$$\begin{aligned}\binom{16}{8} &= \binom{8}{0}^2 + \binom{8}{1}^2 + \binom{8}{2}^2 + \binom{8}{3}^2 + \binom{8}{4}^2 + \binom{8}{5}^2 + \binom{8}{6}^2 + \binom{8}{7}^2 + \binom{8}{8}^2 \\ &= 1^2 + 8^2 + 28^2 + 56^2 + 70^2 + 56^2 + 28^2 + 8^2 + 1^2 \\ &= 12{,}870\end{aligned}$$

▲

Example 1.25 A little girl has six building blocks and is required to select four of them at a time to build a model. If the order of the blocks in each model is not important, how many models can the little girl build?

Solution The number of models is

$$C(6,4) = \frac{6!}{(6-4)!4!} = \frac{6!}{2!4!} = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{2 \times 1 \times 4 \times 3 \times 2 \times 1} = 15$$

▲

Recall that when the order of the blocks is important, we would have $P(6,4) = 360$ models. Also $P(6,4)/C(6,4) = 24 = 4!$, which indicates that for each combination, there are 4! arrangements involved.

Example 1.26 Five boys and five girls are getting together for a party.

a. How many couples can be formed?
b. Suppose one of the boys has two sisters among the five girls, and he would not accept either of them as a partner. How many couples can be formed?

Solution (a) Without any restriction, there are five girls with whom each of the boys can be matched. Thus, the number of couples that can be formed is $5 \times 5 = 25$.

(b) The boy who has two sisters among the girls can only be matched with three girls, but each of the other four boys can be matched with any of the girls. Thus, the number of possible couples is given by $3 + 4 \times 5 = 23$.

▲

1.10.5 The Binomial Theorem

The following theorem, which is called the binomial theorem, is presented without proof. The theorem states that

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

This theorem can be used to present a more formal proof of the statement we made earlier in the chapter about the number of subsets of a set with n elements. The number of subsets of size k is $\binom{n}{k}$. Thus, summing this over all possible values of k we obtain the desired result:

$$\sum_{k=0}^{n} \binom{n}{k} = \sum_{k=0}^{n} \binom{n}{k} 1^k (1)^{n-k} = (1+1)^n = 2^n$$

1.10.6 Stirling's Formula

Problems involving permutations and combinations require the calculation of $n!$. Prior to the advent of the current powerful handheld calculators, the evaluation of $n!$ was tedious even for a moderately large n. Because of this, an approximate formula called the Stirling's formula was developed to obtain values of $n!$. Studies indicate that this formula gives very good results especially for large values of n. The Stirling's formula is given by

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n = \sqrt{2\pi n}\, n^n e^{-n}$$

where $e = 2.71828\ldots$ is the base of the natural logarithms and the notation $a \sim b$ means that the number on the right is an asymptotic representation of the number on the left. As a check on the accuracy of the formula, by direct computation $10! = 3{,}628{,}800$, while the value obtained via the Stirling's formula is 3.60×10^6, which represents an error of 0.79%. In general, the percentage error in the approximation is about $100/12n$.

Example 1.27 Evaluate 50!

Solution Using the Stirling's formula we obtain

$$50! = \sqrt{100\pi}\, 50^{50} e^{-50} = 10\sqrt{\pi}\left(\frac{50}{2.71828}\right)^{50} = 3.04 \times 10^{64}$$

▲

Example 1.28 Evaluate 70!

Solution Using the Stirling's formula we obtain

$$
\begin{aligned}
70! &= \sqrt{140\pi}\,70^{70}e^{-70} = N \\
\log N &= \frac{1}{2}\log 140 + \frac{1}{2}\log \pi + 70\log 70 - 70\log e \\
&= 1.07306 + 0.24857 + 129.15686 - 30.40061 = 100.07788 = 0.07788 + 100 \\
N &= 1.20 \times 10^{100}
\end{aligned}
$$

▲

1.10.7 Applications of Combinations in Probability

As we shall see in Chapter 4, combination plays a very important role in the class of random variables that have the binomial distribution as well as those that have the hypergeometric distribution. In this section, we discuss how it can be applied to the problem of counting the number of selections among items that contain two subgroups. To understand these applications, we first state the following ***fundamental counting rule*** [2]:

> Assume that a number of multiple choices are to be made, which include m_1 ways of making the first choice, m_2 ways of making the second choice, m_3 ways of making the third choice, and so on. If these choices can be made independently, then the total number of possible ways of making these choices is $m_1 \times m_2 \times m_3 \times \cdots$

Example 1.29 The standard car license plate in a certain U.S. state has seven characters that are made up as follows. The first character is one of the digits 1, 2, 3, or 4; the next three characters are letters $(a, b, \ldots, z)$ of which repetition is allowed; and the final three characters are digits $(0, 1, \ldots, 9)$ that also allow repetition.

a. How many license plates are possible?

b. How many of these possible license plates have no repeated characters?

Solution Let m_1 be the number of ways of choosing the first character, m_2 the number of ways of choosing the next three characters, and m_3 the number of ways of choosing the final three characters. Since these choices can be made independently, the principle of the fundamental counting rule implies that there are $m_1 \times m_2 \times m_3$ total number of possible ways of making these choices.

(a) $m_1 = C(4,1) = 4$; since repetition is allowed, $m_2 = \{C(26,1)\}^3 = 26^3$; and since repetition is allowed, $m_3 = \{C(10,1)\}^3 = 10^3$. Thus, the number of possible license plates is $4 \times 26^3 \times 10^3 = 70{,}304{,}000$.

(b) When repetition is not allowed, we obtain $m_1 = C(4,1) = 4$. To obtain the new m_2 we note that after the first letter has been chosen, it cannot be chosen again as the second or third letter, and after the second letter has been chosen, it cannot be chosen as the third letter. This means that there are $m_2 = C(26,1) \times C(25,1) \times C(24,1) = 26 \times 25 \times 24$ ways of choosing the next three letters of the plate. Similarly, since repetition is not allowed, the digit chosen as the first character of the license plate cannot appear in the third set of characters. This means that the first digit of the third set of characters will be chosen from nine digits, the second from eight digits, and the third from seven digits. Thus, we have that $m_3 = 9 \times 8 \times 7$. Therefore, the number of possible license plates that have no repeated characters is given by

$$M = 4 \times 26 \times 25 \times 24 \times 9 \times 8 \times 7 = 31{,}449{,}600$$

▲

Example 1.30 Suppose there are k defective items in a box that contains m items. How many samples of n items of which j items are defective can we get from the box?

Solution Since there are two classes of items (defective versus nondefective), we can select independently from each group once the number of defective items in the sample has been specified. Thus, since there are k defective items in the box, the total number of ways of selecting j out of the k items at a time is $C(k,j) = \binom{k}{j}$, where $0 \le j \le \min(k,n)$. Similarly, since there are $m-k$ nondefective items in the box, the total number of ways of selecting $n-j$ of them at a time is $C(m-k, n-j) = \binom{m-k}{n-j}$. Since these two choices can be made independently, the total number of ways of choosing j defective items and $n-j$ nondefective items is $C(k,j) \times C(m-k, n-j)$, which is

$$C(k,j)C(m-k,n-j) = \binom{k}{j}\binom{m-k}{n-j}$$

▲

Example 1.31 A container has 100 items, 5 of which are defective. If we pick samples of 20 items from the container, find the total number of samples with at most one bad item among them.

Solution Let A be the event that there is no defective item in the selected sample and B the event that there is exactly one defective item in the selected sample. Then event A consists of two subevents: zero defective items and 20 nondefective items. Similarly, event B consists of two subevents: 1 defective item and 19 nondefective items. The number of ways in which event A can occur is $C(5,0) \times C(95,20) = C(95,20)$. Similarly, the number of ways in which event B can occur is $C(5,1) \times C(95,19) = 5C(95,19)$. Therefore, the total number of samples with at most one defective item is the sum of the two, which is

$$C(95,20) + 5C(95,19) = \frac{95!}{75!20!} + \frac{95! \times 5}{76!19!} = \frac{176 \times 95!}{76!20!} = 3.96 \times 10^{20}$$

Note that when we use the Stirling's formula we get

$$C(95,20) + 5C(95,19) = \frac{176 \times 95!}{76!20!} = \frac{176 \times 95^{95.5}}{20^{20.5} \times 76^{76.5} \times \sqrt{2\pi}} = K$$

$$\begin{aligned} \log K &= \log 176 + 95.5 \log 95 - 20.5 \log 20 \\ &\quad - 76.5 \log 76 - 0.5 \log 2 - 0.5 \log \pi \\ &= 20.165772 = 0.165772 + 20 \\ K &= 1.46 \times 10^{20} \end{aligned}$$

which is far less than the correct result. ▲

Example 1.32 A particular department of a small college has seven faculty members of whom two are full professors, three are associate professors, and two are assistant professors. How many committees of three faculty members can be formed if each subgroup (that is, full, associate, and assistant professors) must be represented?

Solution There are $C(2,1) \times C(3,1) \times C(2,1) = 12$ possible committees. ▲

Example 1.33 A batch of 100 manufactured components is checked by an inspector who examines 10 components selected at random. If none of the 10 components is defective, the inspector accepts the whole batch. Otherwise, the batch is subjected to further inspection. What is the probability that a batch containing 10 defective components will be accepted?

Solution Let N denote the number of ways of indiscriminately selecting 10 components from a batch of 100 components. Then N is given by

$$N = C(100, 10) = \frac{100!}{90! \times 10!}$$

Let E denote the event "the batch containing 10 defective components is accepted by the inspector." The number of ways that E can occur is the number of ways of selecting 10 components from the 90 nondefective components and no component from the 10 defective components. This number, $N(E)$, is given by

$$N(E) = C(90, 10) \times C(10, 0) = C(90, 10) = \frac{90!}{80! \times 10!}$$

Because the components are selected at random, the combinations are equiprobable. Thus, the probability of event E is given by

$$\begin{aligned} P(E) &= \frac{N(E)}{N} = \frac{90!}{80! \times 10!} \times \frac{90! \times 10!}{100!} \\ &= \frac{90! \times 90!}{100! \times 80!} = \frac{90 \times 89 \times \cdots \times 81}{100 \times 99 \times \cdots \times 91} \\ &= 0.3305 \end{aligned}$$

▲

Example 1.34 The Applied Probability professor gave the class a set of 12 review problems and told them that the midterm exam would consist of 6 of the 12 problems selected at random. If Lidya memorized the solutions to 8 of the 12 problems but could not solve any of the other 4 problems, what is the probability that she got 4 or more problems correct in the exam?

Solution By choosing to memorize only a subset of the review problems Lidya partitioned the 12 problems into two sets: a set consisting of the 8 problems she memorized and a set consisting of the 4 problems she could not solve. If she got k problems correct in the exam, then the k problems came from the first set and the $6 - k$ problems she failed came from the second set, where $k = 0, 1, 2, \ldots, 6$. The number of ways of choosing 6 problems from 12 problems is $C(12, 6)$. The number of ways of choosing k problems from the 8 problems that she memorized is $C(8, k)$, and the number of ways of choosing $6 - k$ problems from the four she did not memorize is $C(4, 6 - k)$, where $6 - k \leq 4$ or $2 \leq k \leq 6$. Because the problems have been partitioned, the number of ways in which the 8 problems can be chosen so that Lidya could get 4 or more of them correct in the exam is

$$C(8, 4)C(4, 2) + C(8, 5)C(4, 1) + C(8, 6)C(4, 0) = 420 + 224 + 28 = 672$$

Thus, the probability p that she got 4 or more problems correct in the exam is given by

$$p = \frac{C(8,4)C(4,2) + C(8,5)C(4,1) + C(8,6)C(4,0)}{C(12,6)} = \frac{672}{924}$$
$$= \frac{8}{11}$$

▲

1.11 Reliability Applications

As discussed earlier in the chapter, reliability theory is concerned with the duration of the useful life of components and systems of components. That is, it is concerned with determining the probability that a system with possibly many components will be functioning at time t. The components of a system can be arranged in two basic configurations: *series* configuration and *parallel* configuration. A real system consists of a mixture of series and parallel components, which can sometimes be reduced to an equivalent system of series configuration or a system of parallel configuration. Figure 1.9 illustrates the two basic configurations.

A system with a series configuration will function iff all its components are functioning, while a system with parallel configuration will function iff at least

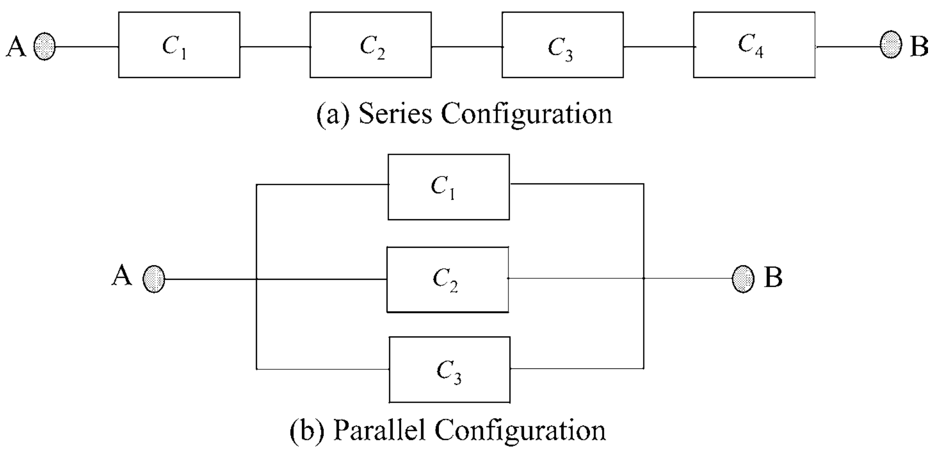

Figure 1.9 Basic Reliability Models

one of the components is functioning. To simplify the discussion, we assume that the different components fail independently.

Consider a system with n components labeled $C_1, C_2, \ldots, C_n$. Let $R_k(t)$ denote the probability that component C_k has not failed in the interval $(0, t]$, where $k = 1, 2, \ldots, n$. That is, $R_k(t)$ is the probability that C_k has not failed up to time t and is called the *reliability function* of C_k. For a system of components in series, the system reliability function is given by

$$R(t) = \prod_{k=1}^{n} R_k(t)$$

This follows from the fact that all components must be operational for the system to be operational.

In the case of a system of parallel components, we need at least one path between A and B for the system to be operational. The probability that no such path exists is the probability that all the components have failed, which is given by $[1 - R_1(t)][1 - R_2(r)] \ldots [1 - R_n(t)]$. Thus, the system reliability function is the complement of this function and is given by

$$R(t) = 1 - [1 - R_1(t)][1 - R_2(t)] \ldots [1 - R_n(t)] = 1 - \prod_{k=1}^{n} [1 - R_k(t)]$$

Example 1.35 Find the system reliability function for the system shown in Figure 1.10 in which C_1 and C_2 are in series and the two are in parallel with C_3.

Solution We first reduce the series structure into a composite component C_4 whose reliability function is given by $R_4(t) = R_1(t)R_2(t)$. Thus, we obtain the new structure shown in Figure 1.11.

Thus, we obtain two parallel components and the system reliability function is

$$\begin{aligned} R(t) &= 1 - [1 - R_3(t)][1 - R_4(t)] \\ &= 1 - [1 - R_3(t)][1 - R_1(t)R_2(t)] \end{aligned}$$

▲

Example 1.36 Find the system reliability function for the system shown in Figure 1.12, which is called a *bridge structure*.

Solution The system is operational if at least one of the following series arrangements is operational: C_1C_4, C_2C_5, $C_1C_3C_5$, or $C_2C_3C_4$. Thus, we can replace

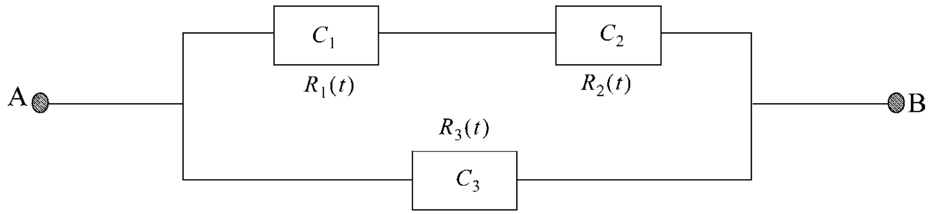

Figure 1.10 Example 1.35

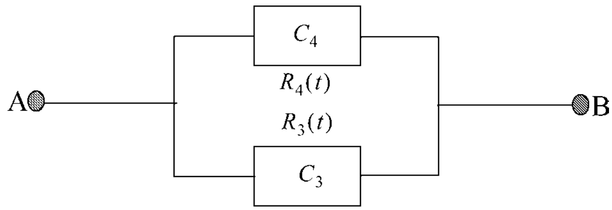

Figure 1.11 Composite System for Example 1.35

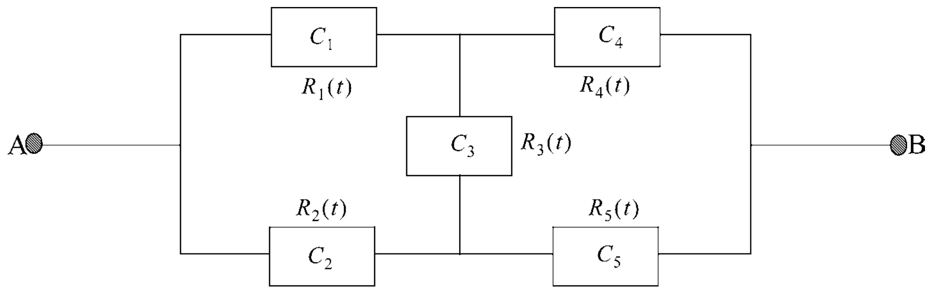

Figure 1.12 Example 1.36

the system with a system of series-parallel arrangements. However, the different paths will not be independent since they have components in common. To avoid this complication, we use a conditional probability approach. First, we consider the reliability function of the system given that C_3 is operational. Next we consider the reliability function of the system given that C_3 is not operational. Figure 1.13 shows the two cases.

When C_3 is operational, the system behaves like a parallel subsystem consisting of C_1 and C_2, which is in series with another parallel subsystem consisting of C_4 and C_5. Thus, if we use shorthand notation and omit the explicit dependence on t, the reliability of the system becomes

$$R_X = [1 - (1 - R_1)(1 - R_2)][1 - (1 - R_4)(1 - R_5)]$$

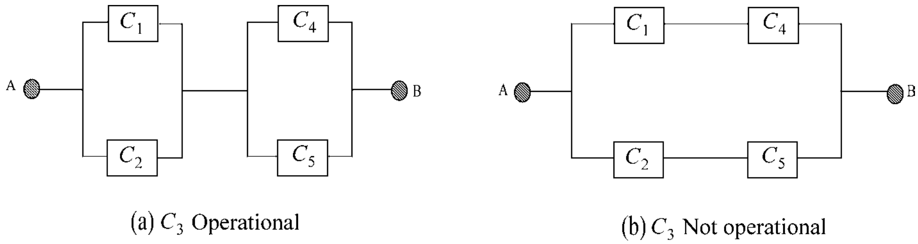

Figure 1.13 Decomposing the System into Two Cases

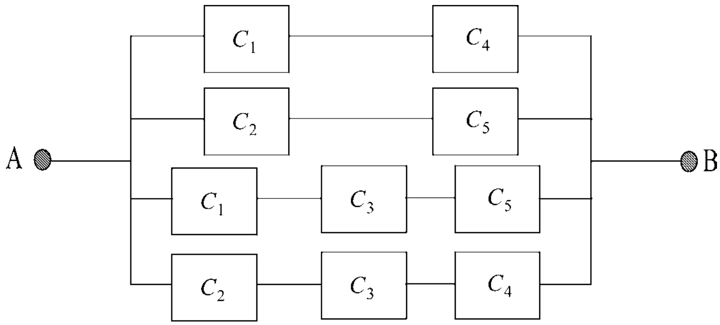

Figure 1.14 Alternative System Configuration for Example 1.36

When C_3 is not operational, signal cannot flow through that component, and the system behaves as shown in Figure 1.13b. Thus, the reliability of the system becomes

$$R_Y = 1 - (1 - R_1R_4)(1 - R_2R_5)$$

Let $P(C_3)$ denote the probability that C_3 is operational in the interval $(0, t]$. Since $P(C_3) = R_3$, we use the law of total probability to obtain the system reliability as follows:

$$\begin{aligned} R &= R_XP(C_3) + R_Y[1 - P(C_3)] = R_XR_3 + R_Y(1 - R_3) \\ &= R_3[1 - (1 - R_1)(1 - R_2)][1 - (1 - R_4)(1 - R_5)] \\ &\quad + (1 - R_3)[1 - (1 - R_1R_4)(1 - R_2R_5)] \\ &= R_1R_4 + R_2R_5 + R_1R_3R_5 + R_2R_3R_4 - R_1R_2R_3R_4 - R_1R_2R_3R_5 \\ &\quad - R_1R_2R_4R_5 - R_1R_3R_4R_5 - R_2R_3R_4R_5 + 2R_1R_2R_3R_4R_5 \end{aligned}$$

The first four positive terms represent the different ways we can pass signals between the input and output. Thus, the equivalent system configuration is as shown in Figure 1.14. The other terms account for the dependencies we mentioned earlier. ▲

Example 1.37 Consider the network shown in Figure 1.15 that interconnects nodes A and B. The switches S_1, S_2, S_3, and S_4 have availabilities A_1, A_2, A_3 and A_4, respectively. That is, the probability that switch S_i is operational at any given time is A_i, $i = 1, 2, 3, 4$. If the switches fail independently, what is the probability that at a randomly selected time A can communicate with B (that is, at least one path can be established between A and B)?

Solution We begin by reducing the structure as shown in Figure 1.16, where S_{1-2} is the composite system of S_1 and S_2, and S_{1-2-3} is the composite system of S_{1-2} and S_3.

From Figure 1.16a, the availability of S_{1-2} is $A_{1-2} = 1-(1-A_1)(1-A_2)$. Similarly, the availability of S_{1-2-3} is $A_{1-2-3} = A_{1-2} \times A_3$. Finally, from Figure 1.16b, the probability that a path exists between A and B is given by

$$\begin{aligned} P_{A-B} &= 1-(1-A_{1-2-3})(1-A_4) \\ &= 1-(1-[1-(1-A_1)(1-A_2)]A_3)(1-A_4) \end{aligned}$$

▲

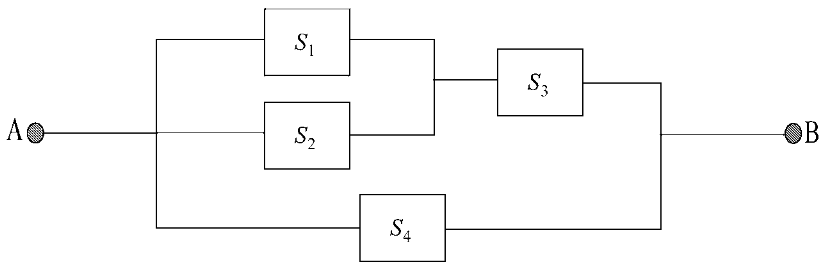

Figure 1.15 Figure for Example 1.37

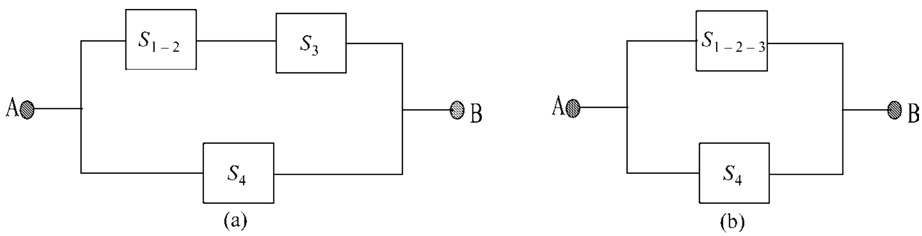

Figure 1.16 Reduced Forms of Figure 1.15

1.12 Chapter Summary

This chapter has developed the basic concepts of probability, random experiments, and events. Several examples are solved and applications of probability have been provided in the fields of communications and reliability engineering. Finally, it introduced the concepts of permutation and combination that will be used in later chapters.

1.13 Problems

Section 1.2: Sample Space and Events

1.1 A fair die is rolled twice. Find the probability of the following events:

a. The second number is twice the first.

b. The second number is not greater than the first.

c. At least one number is greater than 3.

1.2 Two distinct dice A and B are rolled. What is the probability of each of the following events?

a. At least one 4 appears.

b. Just one 4 appears.

c. The sum of the face values is 7.

d. One of the values is 3 and the sum of the two values is 5.

e. One of the values is 3 or the sum of the two values is 5.

1.3 Consider an experiment that consists of rolling a die twice.

a. Plot the sample space S of the experiment.

b. Identify the event A, which is the event that the sum of the two outcomes is equal to 6.

c. Identify the event B, which is the event that the difference between the two outcomes is equal to 2.

1.4 A four-sided fair die is rolled twice. What is the probability that the outcome of the first roll is greater than the outcome of the second roll?

1.5 A coin is tossed until the first head appears, and then the experiment is stopped. Define a sample space for the experiment.

1.6 A coin is tossed four times and observed to be either a head or a tail each time. Describe the sample space for the experiment.

1.7 Three friends, Bob, Chuck, and Dan take turns (in that order) throwing a die until the first "six" appears. The person that throws the first six wins the game, and the game ends. Write down a sample space for this game.

Section 1.3: Definitions of Probability

1.8 A small country has a population of 17 million people of whom 8.4 million are male and 8.6 million are female. If 75% of the male population and 63% of the female population are literate, what percentage of the total population is literate?

1.9 Let A and B be two independent events with $P[A]=0.4$ and $P[A\cup B]=0.7$. What is $P[B]$?

1.10 Consider two events A and B with known probabilities $P[A]$, $P[B]$, and $P[A\cap B]$. Find the expression for the event that exactly one of the two events occurs in terms of $P[A]$, $P[B]$, and $P[A\cap B]$.

1.11 Two events A and B have the following probabilities: $P[A]=1/4$, $P[B|A]=1/2$, and $P[A|B]=1/3$. Compute (a) $P[A\cap B]$, (b) $P[B]$, and (c) $P[A\cup B]$.

1.12 Two events A and B have the following probabilities: $P[A]=0.6$, $P[B]=0.7$, and $P[A\cap B]=p$. Find the range of values that p can take.

1.13 Two events A and B have the following probabilities: $P[A]=0.5$, $P[B]=0.6$, and $P[\overline{A}\cap\overline{B}]=0.25$. Find the value of $P[A\cap B]$.

1.14 Two events A and B have the following probabilities: $P[A]=0.4$, $P[B]=0.5$, and $P[A\cap B]=0.3$. Calculate the following:

a. $P[A\cup B]$

b. $P[A\cap\overline{B}]$

c. $P[\overline{A}\cup\overline{B}]$

1.15 Christie is taking a multiple-choice test in which each question has four possible answers. She knows the answers to 40% of the questions and can narrow the choices down to two answers 40% of the time. If she knows nothing about the remaining 20% of the questions, what is the probability that she will correctly answer a question chosen at random from the test?

1.16 A box contains nine red balls, six white balls, and five blue balls. If three balls are drawn successively from the box, determine the following:

a. The probability that they are drawn in the order red, white, and blue if each ball is replaced after it has been drawn.

b. The probability that they are drawn in the order red, white, and blue if each ball is not replaced after it has been drawn.

1.17 Let A be the set of positive even integers, let B be the set of positive integers that are divisible by 3, and let C be the set of positive odd integers. Describe the following events:

a. $E_1 = A \cup B$

b. $E_2 = A \cap B$

c. $E_3 = A \cap C$

d. $E_4 = (A \cup B) \cap C$

e. $E_5 = A \cup (B \cap C)$

1.18 A box contains four red balls labeled R_1, R_2, R_3, and R_4; and three white balls labeled W_1, W_2, and W_3. A random experiment consists of drawing a ball from the box. State the outcomes of the following events:

a. E_1, the event that the number on the ball (i.e., the subscript of the ball) is even.

b. E_2, the event that the color of the ball is red and its number is greater than 2.

c. E_3, the event that the number on the ball is less than 3.

d. $E_4 = E_1 \cup E_3$.

e. $E_5 = E_1 \cup (E_2 \cap E_3)$.

1.19 A box contains 50 computer chips of which 8 are known to be bad. A chip is selected at random and tested.

(a) What is the probability that it is bad?

(b) If a test on the first chip shows that it is bad, what is the probability that a second chip selected at random will also be bad, assuming the tested chip is not put back into the box?

(c) If the first chip tests good, what is the probability that a second chip selected at random will be bad, assuming the tested chip is not put back into the box?

Section 1.5: Elementary Set Theory

1.20 A set S has four members: A, B, C, and D. Determine all possible subsets of S.

1.21 For three sets A, B, and C, use the Venn diagram to show the areas corresponding to the sets (a) $(A \cup C) - C$, (b) $\overline{B} \cap A$, (c) $A \cap B \cap C$, and (d) $(\overline{A \cup B}) \cap C$.

1.22 A universal set is given by $S = \{2, 4, 6, 8, 10, 12, 14\}$. If we define two sets $A = \{2, 4, 8\}$ and $B = \{4, 6, 8, 12\}$, determine the following: (a) $\overline{A}$, (b) $B - A$, (c) $A \cup B$, (d) $A \cap B$, (e) $\overline{A} \cap B$, and (f) $(A \cap B) \cup (\overline{A} \cap B)$.

1.23 Consider the switching networks shown in Figure 1.17. Let E_k denote the event that switch S_k is closed, $k = 1, 2, 3, 4$. Let E_{AB} denote the event that there is a closed path between nodes A and B. Express E_{AB} in terms of the E_k for each network.

1.24 Let A, B, and C be three events. Write out the expressions for the following events in terms of A, B, and C using set notation:

a. A occurs but neither B nor C occurs.

b. A and B occur, but not C.

c. A or B occurs, but not C.

d. Either A occurs and not B, or B occurs and not A.

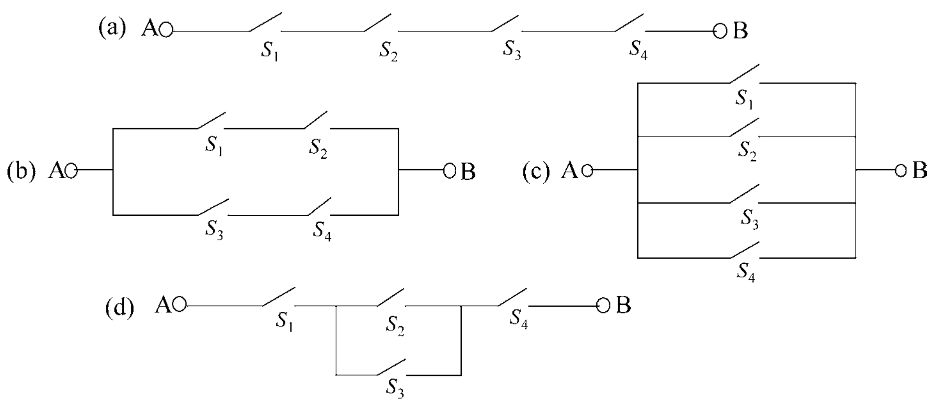

Figure 1.17 Figure for Problem 1.23

Section 1.6: Properties of Probability

1.25 Mark and Lisa registered for Physics 101 class. Mark attends class 65% of the time and Lisa attends class 75% of the time. Their absences are independent. On a given day, what is the probability that

(a) at least one of them is in class?

(b) exactly one of them is in class?

(c) Mark is in class, given that only one of them is in class?

1.26 The probability of rain on a day of the year selected at random is 0.25 in a certain city. The local weather forecast is correct 60% of the time when the forecast is rain and 80% of the time for other forecasts. What is the probability that the forecast on a day selected at random is correct?

1.27 53% of the adults in a certain city are female, and 15% of the adults are unemployed males.

(a) What is the probability that an adult chosen at random in this city is an employed male?

(b) If the overall unemployment rate in the city is 22%, what is the probability that a randomly selected adult is an employed female?

1.28 A survey of 100 companies shows that 75 of them have installed wireless local area networks (WLANs) on their premises. If three of these companies are chosen at random without replacement, what is the probability that each of the three has installed WLANs?

Section 1.7: Conditional Probability

1.29 A certain manufacturer produces cars at two factories labeled A and B. Ten percent of the cars produced at factory A are found to be defective, while 5% of the cars produced at factory B are defective. If factory A produces 100,000 cars per year and factory B produces 50,000 cars per year, compute the following:

(a) The probability of purchasing a defective car from the manufacturer

(b) If a car purchased from the manufacturer is defective, what is the probability that it came from factory A?

1.30 Kevin rolls two dice and tells you that there is at least one 6. What is the probability that the sum is at least 9?

1.31 Chuck is a fool with probability 0.6, a thief with probability 0.7, and neither with probability 0.25.

(a) What is the probability that he is a fool or a thief but not both?

(b) What is the conditional probability that he is a thief, given that he is not a fool?

1.32 Studies indicate that the probability that a married man votes is 0.45, the probability that a married woman votes is 0.40, and the probability that a married woman votes given that her husband does is 0.60. Compute the following probabilities:

(a) Both a man and his wife vote.

(b) A man votes given that his wife does.

1.33 Tom is planning to pick up a friend at the airport. He has figured out that the plane is late 80% of the time when it rains, but only 30% of the time when it does not rain. If the weather forecast that morning calls for a 40% chance of rain, what is the probability that the plane will be late?

1.34 Consider the communication channel shown in Figure 1.18. The symbols transmitted are 0 and 1. However, three possible symbols can be received: 0, 1, and E. Thus, we define the input symbol set as $X \in \{0, 1\}$ and the output symbol set as $Y \in \{0, 1, E\}$. The transition (or conditional) probabilities are defined by $p_{Y|X}$, which is the probability that Y is received, given that X was transmitted. In particular, $p_{0|0} = 0.8$ (i.e., given that 0 is transmitted, it is received as 0 with probability 0.8), $p_{1|0} = 0.1$ (i.e., given that 0 is transmitted, it is received as 1 with probability 0.1), and $p_{E|0} = 0.1$ (i.e., given that 0 is transmitted, it is received as E with probability 0.1). Similarly, $p_{0|1} = 0.2$, $p_{1|1} = 0.7$, and $p_{E|1} = 0.1$. If $P[X = 0] = P[X = 1] = 0.5$, determine the following:

(a) $P[Y = 0]$, $P[Y = 1]$, and $P[Y = E]$

(b) If 0 is received, what is the probability that 0 was transmitted?

(c) If E is received, what is the probability that 1 was transmitted?

(d) If 1 is received, what is the probability that 1 was transmitted?

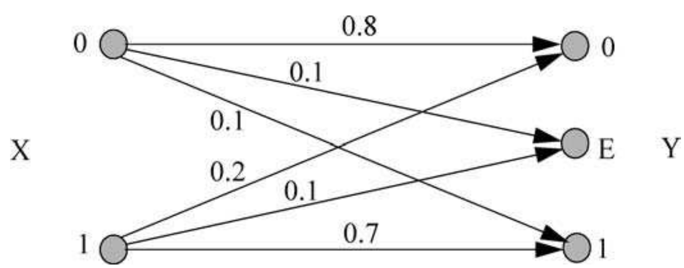

Figure 1.18 Figure for Problem 1.34

1.35 A group of students consists of 60% men and 40% women. Among the men, 30% are foreign students, and among the women, 20% are foreign students, A student is randomly selected from the group and found to be a foreign student. What is the probability that the student is a woman?

1.36 Joe frequently gets into trouble at school, and past experience shows that 80% of the time he is guilty of the offense he is accused of. Joe has just gotten into trouble again, and two other students, Chris and Dana, have been called into the principal's office to testify about the incident. Chris is Joe's friend and will tell the truth if Joe is innocent, but will lie with probability 0.2 if Joe is guilty. Dana does not like Joe and so will tell the truth if Joe is guilty, but will lie with probability 0.3 if Joe is innocent.

a. What is the probability that Chris and Dana give conflicting testimonies?

b. What is the probability that Joe is guilty, given that Chris and Dana give conflicting testimonies?

1.37 Three car brands A, B, and C, have all the market share in a certain city. Brand A has 20% of the market share, brand B has 30%, and brand C has 50%. The probability that a brand A car needs a major repair during the first year of purchase is 0.05, the probability that a brand B car needs a major repair during the first year of purchase is 0.10, and the probability that a brand C car needs a major repair during the first year of purchase is 0.15.

a. What is the probability that a randomly selected car in the city needs a major repair during its first year of purchase?

b. If a car in the city needs a major repair during its first year of purchase, what is the probability that it is a brand A car?

Section 1.8: Independent Events

1.38 If I toss two coins and tell you that at least one is heads, what is the probability that the first coin is heads?

1.39 Assume that we roll two dice and define three events A, B, and C, where $A = \{\text{The first die is odd}\}$, $B = \{\text{The second die is odd}\}$, and $C = \{\text{The sum is odd}\}$. Show that these events are pairwise independent but the three are not independent.

1.40 Consider a game that consists of two successive trials. The first trial has outcome A or B, and the second trial has outcome C or D. The probabilities of the four possible outcomes of the game are as follows:

Outcome	AC	AD	BC	BD
Probability	1/3	1/6	1/6	1/3

Determine in a convincing way if A and C are statistically independent.

1.41 Suppose that two events A and B are mutually exclusive and $P[B] > 0$. Under what conditions will A and B be independent?

Section 1.10: Combinatorial Analysis

1.42 Four married couples bought tickets for eight seats in a row for a football game.

a. In how many different ways can they be seated?

b. In how many ways can they be seated if each couple is to sit together with the husband to the left of his wife?

c. In how many ways can they be seated if each couple is to sit together?

d. In how many ways can they be seated if all the men are to sit together and all the women are to sit together?

1.43 A committee consisting of three electrical engineers and three mechanical engineers is to be formed from a group of seven electrical engineers and five mechanical engineers. Find the number of ways in which this can be done if

a. any electrical engineer and any mechanical engineer can be included.

b. one particular electrical engineer must be on the committee.

c. two particular mechanical engineers cannot be on the same committee.

1.44 Use Stirling's formula to evaluate 200!

1.45 A committee of three members is to be formed consisting of one representative from labor, one from management, and one from the public. If there are seven possible representatives from labor, four from management, and five from the public, how many different committees can be formed?

1.46 There are 100 U.S. senators, two from each of the 50 states.

(a) If two senators are chosen at random, what is the probability that they are from the same state?

(b) If ten senators are randomly chosen to form a committee, what is the probability that they are all from different states?

1.47 A committee of seven people is to be formed from a pool of 10 men and 12 women.

(a) What is the probability that the committee will consist of three men and four women?

(b) What is the probability that the committee will consist of all men?

1.48 Five departments in the college of engineering, which are labeled departments A, B, C, D, and E, send three delegates each to the college's convention. A committee of four delegates, selected by lot, is formed. Determine the probability that

(a) Department A is not represented on the committee.

(b) Department A has exactly one representative on the committee.

(c) Neither department A nor department C is represented on the committee.

Section 1.11: Reliability Applications

1.49 Consider the system shown in Figure 1.19. If the number inside each box indicates the probability that the component will independently fail within the next two years, find the probability that the system fails within two years.

1.50 Consider the structure shown in Figure 1.20. Switches S_1 and S_2 are in series, and the pair is in parallel with a parallel arrangement of switches S_3

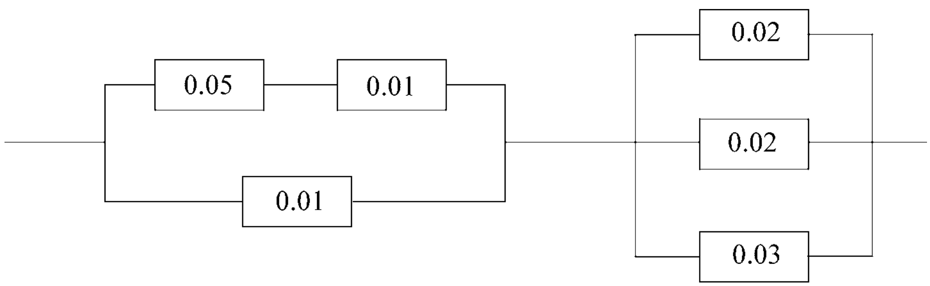

Figure 1.19 Figure for Problem 1.49

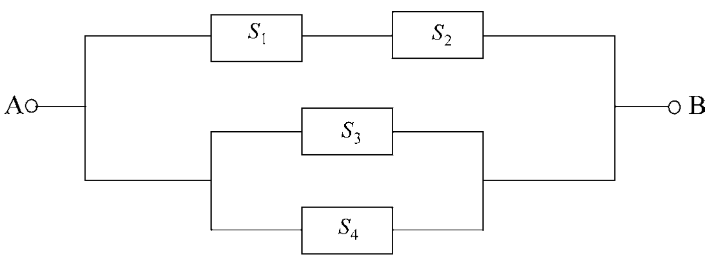

Figure 1.20 Figure for Problem 1.50

and S_4. Their reliability functions are $R_1(t)$, $R_2(t)$, $R_3(t)$, and $R_4(t)$, respectively. The structure interconnects nodes A and B. What is the reliability function of the composite system in terms of $R_1(t)$, $R_2(t)$, $R_3(t)$, and $R_4(t)$ if the switches fail independently?

1.51 Consider the network shown in Figure 1.21 that interconnects nodes A and B. The switches labeled $S_1, S_2, \ldots, S_8$ have the reliability functions $R_1(t), R_2(t), \ldots, R_8(t)$, respectively. If the switches fail independently, find the reliability function of the composite system.

1.52 Consider the network shown in Figure 1.22 that interconnects nodes A and B. The switches labeled $S_1, S_2, \ldots, S_8$ have the reliability functions $R_1(t), R_2(t), \ldots, R_8(t)$, respectively. If the switches fail independently, find the reliability function of the composite system.

1.53 Consider the network shown in Figure 1.23 that interconnects nodes A and B. The switches labeled $S_1, S_2, \ldots, S_7$ have the reliability functions $R_1(t), R_2(t), \ldots, R_7(t)$, respectively. If the switches fail independently, find the reliability function of the composite system.

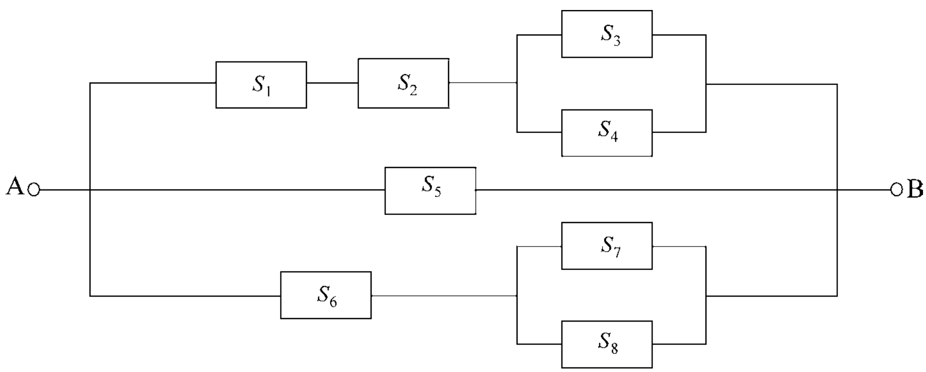

Figure 1.21 Figure for Problem 1.51

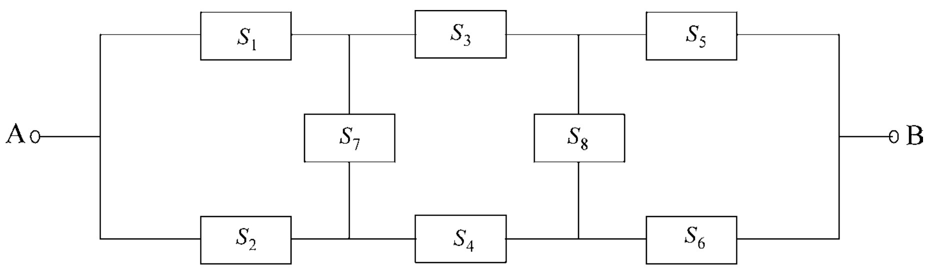

Figure 1.22 Figure for Problem 1.52

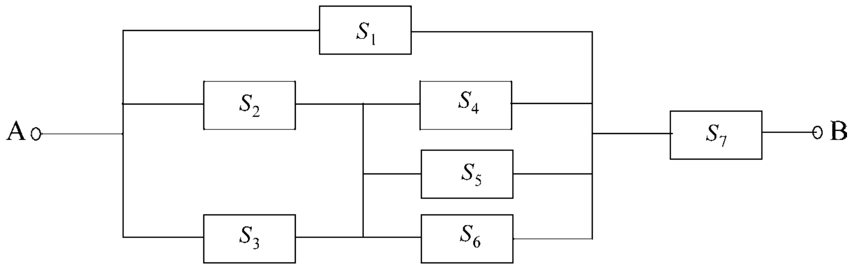

Figure 1.23 Figure for Problem 1.53

1.14 References

The following books provide basic information on probability at the same level presented in this chapter.

[1] Ash, C., *The Probability Tutoring Book*, IEEE Press, New York, 1993.
[2] Chung, K.L., *Elementary Probability Theory with Stochastic Processes*, 3rd ed., Springer-Verlag, New York, 1979.
[3] Clarke, A.B., and Disney, R.L., *Probability and Random Processes: A First Course with Applications*, John Wiley & Sons, New York, 1985.
[4] Drake, A.W., *Fundamentals of Applied Probability Theory*, McGraw-Hill, New York, 1967.
[5] Freund, J.E., *Introduction to Probability*, Dickenson Publishing Company, Encino, CA, 1973. Reprinted by Dover Publications, Inc., New York, 1993.
[6] Goldberg, S., *Probability: An Introduction*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1960. Reprinted by Dover Publications, Inc., New York.
[7] Haigh, J., *Probability Models*, Springer-Verlag, London, 2002.
[8] Parzen, E., *Modern Probability Theory and Its Applications*, John Wiley & Sons, New York, 1960.
[9] Pfeiffer, P.E., *Concepts of Probability Theory*, McGraw-Hill, New York, 1965. Reprinted by Dover Publications, Inc., New York, in 1978.
[10] Ross, S., *A First Course in Probability*, 6th ed., Prentice-Hall, Upper Saddle River, New Jersey, 2002.
[11] Rozanov, Y.A., *Probability Theory: A Concise Course*, Dover Publications, New York, 1977.
[12] Thomas, J.B., *An Introduction to Applied Probability and Random Processes*, Robert E. Krieger Publishing Company, Malabar, Florida, 1981.
[13] Trivedi, K.S., *Probability and Statistics with Reliability, Queueing and Computer Science Applications*, 2nd ed., John Wiley & Sons, New York, 2002.
[14] Tuckwell, H.C., *Elementary Applications of Probability Theory*, 2nd ed., Chapman & Hall, London, 1995.