Chapter 1

Probability

1.1 INTRODUCTION

The study of probability stems from the analysis of certain games of chance, and it has found applications in most branches of science and engineering. In this chapter the basic concepts of probability theory are presented.

1.2 SAMPLE SPACE AND EVENTS

A. Random Experiments:

In the study of probability, any process of observation is referred to as an *experiment*. The results of an observation are called the *outcomes* of the experiment. An experiment is called a *random experiment* if its outcome cannot be predicted. Typical examples of a random experiment are the roll of a die, the toss of a coin, drawing a card from a deck, or selecting a message signal for transmission from several messages.

B. Sample Space:

The set of all possible outcomes of a random experiment is called the *sample space* (or *universal set*), and it is denoted by S. An element in S is called a *sample point*. Each outcome of a random experiment corresponds to a sample point.

EXAMPLE 1.1 Find the sample space for the experiment of tossing a coin (*a*) once and (*b*) twice.

(*a*) There are two possible outcomes, heads or tails. Thus

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

where H and T represent head and tail, respectively.

(*b*) There are four possible outcomes. They are pairs of heads and tails. Thus

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

EXAMPLE 1.2 Find the sample space for the experiment of tossing a coin repeatedly and of counting the number of tosses required until the first head appears.

Clearly all possible outcomes for this experiment are the terms of the sequence 1, 2, 3, ... Thus

$$S = \{1, 2, 3, \ldots\}$$

Note that there are an infinite number of outcomes.

EXAMPLE 1.3 Find the sample space for the experiment of measuring (in hours) the lifetime of a transistor.

Clearly all possible outcomes are all nonnegative real numbers. That is,

$$S = \{\tau: 0 \le \tau \le \infty\}$$

where τ represents the life of a transistor in hours.

Note that any particular experiment can often have many different sample spaces depending on the observation of interest (Probs. 1.1 and 1.2). A sample space S is said to be *discrete* if it consists of a finite number of sample points (as in Example 1.1) or countably infinite sample points (as in Example 1.2). A set is called *countable* if its elements can be placed in a one-to-one correspondence with the positive integers. A sample space S is said to be *continuous* if the sample points constitute a continuum (as in Example 1.3).

C. Events:

Since we have identified a sample space S as the set of all possible outcomes of a random experiment, we will review some set notations in the following.

If ζ is an element of S (or belongs to S), then we write

$$\zeta \in S$$

If S is not an element of S (or does not belong to S), then we write

$$\zeta \notin S$$

A set A is called a *subset* of B, denoted by

$$A \subset B$$

if every element of A is also an element of B. Any subset of the sample space S is called an *event*. A sample point of S is often referred to as an *elementary event*. Note that the sample space S is the subset of itself, that is, $S \subset S$. Since S is the set of all possible outcomes, it is often called the *certain event*.

EXAMPLE 1.4 Consider the experiment of Example 1.2. Let A be the event that the number of tosses required until the first head appears is even. Let B be the event that the number of tosses required until the first head appears is odd. Let C be the event that the number of tosses required until the first head appears is less than 5. Express events A, B, and C.

$$A = \{2, 4, 6, \ldots\}$$
$$B = \{1, 3, 5, \ldots\}$$
$$C = \{1, 2, 3, 4\}$$

1.3 ALGEBRA OF SETS

A. Set Operations:

1. Equality:

Two sets A and B are equal, denoted $A = B$, if and only if $A \subset B$ and $B \subset A$.

2. Complementation:

Suppose $A \subset S$. The *complement* of set A, denoted $\bar{A}$, is the set containing all elements in S but not in A.

$$\bar{A} = \{\zeta : \zeta \in S \text{ and } \zeta \notin A\}$$

3. Union:

The *union* of sets A and B, denoted $A \cup B$, is the set containing all elements in either A or B or both.

$$A \cup B = \{\zeta : \zeta \in A \text{ or } \zeta \in B\}$$

4. Intersection:

The *intersection* of sets A and B, denoted $A \cap B$, is the set containing all elements in both A and B.

$$A \cap B = \{\zeta : \zeta \in A \text{ and } \zeta \in B\}$$

5. Null Set:

The set containing no element is called the *null set*, denoted $\varnothing$. Note that

$$\varnothing = \bar{S}$$

6. Disjoint Sets:

Two sets A and B are called *disjoint* or *mutually exclusive* if they contain no common element, that is, if $A \cap B = \varnothing$.

The definitions of the union and intersection of two sets can be extended to any finite number of sets as follows:

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n$$
$$= \{\zeta: \zeta \in A_1 \text{ or } \zeta \in A_2 \text{ or } \cdots \zeta \in A_n\}$$

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \cdots \cap A_n$$
$$= \{\zeta: \zeta \in A_1 \text{ and } \zeta \in A_2 \text{ and } \cdots \zeta \in A_n\}$$

Note that these definitions can be extended to an infinite number of sets:

$$\bigcup_{i=1}^{\infty} A_i = A_1 \cup A_2 \cup A_3 \cup \cdots$$
$$\bigcap_{i=1}^{\infty} A_i = A_1 \cap A_2 \cap A_3 \cap \cdots$$

In our definition of event, we state that every subset of S is an event, including S and the null set $\varnothing$. Then

$$S = \text{the certain event}$$
$$\varnothing = \text{the impossible event}$$

If A and B are events in S, then

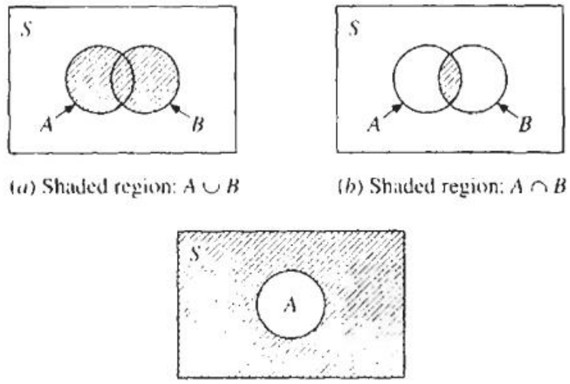
$$\bar{A} = \text{the event that } A \text{ did not occur}$$
$$A \cup B = \text{the event that either } A \text{ or } B \text{ or both occurred}$$
$$A \cap B = \text{the event that both } A \text{ and } B \text{ occurred}$$

Similarly, if $A_1, A_2, \ldots, A_n$ are a sequence of events in S, then

$$\bigcup_{i=1}^{n} A_i = \text{the event that at least one of the } A_i \text{ occurred;}$$
$$\bigcap_{i=1}^{n} A_i = \text{the event that all of the } A_i \text{ occurred.}$$

B. Venn Diagram:

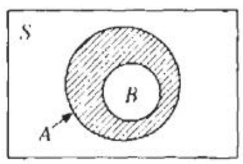
A graphical representation that is very useful for illustrating set operation is the Venn diagram. For instance, in the three Venn diagrams shown in Fig. 1-1, the shaded areas represent, respectively, the events $A \cup B$, $A \cap B$, and $\bar{A}$. The Venn diagram in Fig. 1-2 indicates that $B \subset A$ and the event $A \cap \bar{B}$ is shown as the shaded area.



(a) Shaded region: $A \cup B$ (b) Shaded region: $A \cap B$

(c) Shaded region: $\bar{A}$

Fig. 1-1



$B \subset A$

Shaded region: $A \cap \bar{B}$

Fig. 1-2

C. Identities:

By the above set definitions or reference to Fig. 1-1, we obtain the following identities:

$$\bar{S} = \varnothing \tag{1.1}$$

$$\bar{\varnothing} = S \tag{1.2}$$

$$\bar{\bar{A}} = A \tag{1.3}$$

$$S \cup A = S \tag{1.4}$$

$$S \cap A = A \tag{1.5}$$

$$A \cup \bar{A} = S \tag{1.6}$$

$$A \cap \bar{A} = \varnothing \tag{1.7}$$

The union and intersection operations also satisfy the following laws:

Commutative Laws:

$$A \cup B = B \cup A \tag{1.8}$$

$$A \cap B = B \cap A \tag{1.9}$$

Associative Laws:

$$A \cup (B \cup C) = (A \cup B) \cup C \tag{1.10}$$

$$A \cap (B \cap C) = (A \cap B) \cap C \tag{1.11}$$

Distributive Laws:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \tag{1.12}$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \tag{1.13}$$

De Morgan's Laws:

$$\overline{A \cup B} = \bar{A} \cap \bar{B} \tag{1.14}$$

$$\overline{A \cap B} = \bar{A} \cup \bar{B} \tag{1.15}$$

These relations are verified by showing that any element that is contained in the set on the left side of the equality sign is also contained in the set on the right side, and vice versa. One way of showing this is by means of a Venn diagram (Prob. 1.13). The distributive laws can be extended as follows:

$$A \cap \left(\bigcup_{i=1}^{n} B_i \right) = \bigcup_{i=1}^{n} (A \cap B_i) \tag{1.16}$$

$$A \cup \left(\bigcap_{i=1}^{n} B_i \right) = \bigcap_{i=1}^{n} (A \cup B_i) \tag{1.17}$$

Similarly, De Morgan's laws also can be extended as follows (Prob. 1.17):

$$\overline{\left(\bigcup_{i=1}^{n} A_i \right)} = \bigcap_{i=1}^{n} \bar{A}_i \tag{1.18}$$

$$\overline{\left(\bigcup_{i=1}^{n} A_i \right)} = \bigcap_{i=1}^{n} \bar{A}_i \tag{1.19}$$

1.4 THE NOTION AND AXIOMS OF PROBABILITY

An assignment of real numbers to the events defined in a sample space S is known as the *probability measure*. Consider a random experiment with a sample space S, and let A be a particular event defined in S.

A. Relative Frequency Definition:

Suppose that the random experiment is repeated n times. If event A occurs $n(A)$ times, then the probability of event A, denoted $P(A)$, is defined as

$$P(A) = \lim_{n \to \infty} \frac{n(A)}{n} \tag{1.20}$$

where $n(A)/n$ is called the relative frequency of event A. Note that this limit may not exist, and in addition, there are many situations in which the concepts of repeatability may not be valid. It is clear that for any event A, the relative frequency of A will have the following properties:

1. $0 \le n(A)/n \le 1$, where $n(A)/n = 0$ if A occurs in none of the n repeated trials and $n(A)/n = 1$ if A occurs in all of the n repeated trials.
2. If A and B are mutually exclusive events, then

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

and

$$\frac{n(A \cup B)}{n} = \frac{n(A)}{n} + \frac{n(B)}{n}$$

B. Axiomatic Definition:

Let S be a finite sample space and A be an event in S. Then in the *axiomatic* definition, the probability $P(A)$ of the event A is a real number assigned to A which satisfies the following three *axioms*:

Axiom 1: $P(A) \geq 0$ (1.21)

Axiom 2: $P(S) = 1$ (1.22)

Axiom 3: $P(A \cup B) = P(A) + P(B) \quad \text{if } A \cap B = \varnothing$ (1.23)

If the sample space S is not finite, then axiom 3 must be modified as follows:

Axiom 3′: If $A_1, A_2, \ldots$ is an infinite sequence of mutually exclusive events in S ($A_i \cap A_j = \varnothing$ for $i \neq j$), then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \tag{1.24}$$

These axioms satisfy our intuitive notion of probability measure obtained from the notion of relative frequency.

C. Elementary Properties of Probability:

By using the above axioms, the following useful properties of probability can be obtained:

1. $P(\bar{A}) = 1 - P(A)$ (1.25)
2. $P(\varnothing) = 0$ (1.26)
3. $P(A) \leq P(B) \quad \text{if } A \subset B$ (1.27)
4. $P(A) \leq 1$ (1.28)
5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (1.29)
6. If $A_1, A_2, \ldots, A_n$ are n arbitrary events in S, then

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i \neq j} P(A_i \cap A_j) + \sum_{i \neq j \neq k} P(A_i \cap A_j \cap A_k) - \cdots (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n) \tag{1.30}$$

where the sum of the second term is over all distinct pairs of events, that of the third term is over all distinct triples of events, and so forth.

7. If $A_1, A_2, \ldots, A_n$ is a finite sequence of mutually exclusive events in S ($A_i \cap A_j = \varnothing$ for $i \neq j$), then

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) \tag{1.31}$$

and a similar equality holds for any subcollection of the events.

Note that property 4 can be easily derived from axiom 2 and property 3. Since $A \subset S$, we have

$$P(A) \leq P(S) = 1$$

Thus, combining with axiom 1, we obtain

$$0 \le P(A) \le 1 \tag{1.32}$$

Property 5 implies that

$$P(A \cup B) \le P(A) + P(B) \tag{1.33}$$

since $P(A \cap B) \ge 0$ by axiom 1.

1.5 EQUALLY LIKELY EVENTS

A. Finite Sample Space:

Consider a finite sample space S with n finite elements

$$S = \{\zeta_1, \zeta_2, \ldots, \zeta_n\}$$

where ζ_i's are elementary events. Let $P(\zeta_i) = p_i$. Then

1. $0 \le p_i \le 1 \qquad i = 1, 2, \ldots, n$ (1.34)
2. $\sum_{i=1}^{n} p_i = p_1 + p_2 + \cdots + p_n = 1$ (1.35)
3. If $A = \bigcup_{i \in I} \zeta_i$, where I is a collection of subscripts, then

$$P(A) = \sum_{\zeta_i \in A} P(\zeta_i) = \sum_{i \in I} p_i \tag{1.36}$$

B. Equally Likely Events:

When all elementary events ζ_i ($i = 1, 2, \ldots, n$) are equally likely, that is,

$$p_1 = p_2 = \cdots = p_n$$

then from Eq. (1.35), we have

$$p_i = \frac{1}{n} \qquad i = 1, 2, \ldots, n \tag{1.37}$$

and

$$P(A) = \frac{n(A)}{n} \tag{1.38}$$

where $n(A)$ is the number of outcomes belonging to event A and n is the number of sample points in S.

1.6 CONDITIONAL PROBABILITY

A. Definition:

The *conditional probability* of an event A given event B, denoted by $P(A \mid B)$, is defined as

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \qquad P(B) > 0 \tag{1.39}$$

where $P(A \cap B)$ is the joint probability of A and B. Similarly,

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \qquad P(A) > 0 \tag{1.40}$$

is the conditional probability of an event B given event A. From Eqs. (*1.39*) and (*1.40*), we have

$$P(A \cap B) = P(A \mid B)P(B) = P(B \mid A)P(A) \tag{1.41}$$

Equation (*1.41*) is often quite useful in computing the joint probability of events.

B. Bayes' Rule:

From Eq. (*1.41*) we can obtain the following *Bayes' rule*:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} \tag{1.42}$$

1.7 TOTAL PROBABILITY

The events $A_1, A_2, \ldots, A_n$ are called *mutually exclusive and exhaustive* if

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n = S \qquad \text{and} \qquad A_i \cap A_j = \varnothing \qquad i \neq j \tag{1.43}$$

Let B be any event in S. Then

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B \mid A_i)P(A_i) \tag{1.44}$$

which is known as the *total probability* of event B (Prob. 1.47). Let $A = A_i$ in Eq. (*1.42*); then, using Eq. (*1.44*), we obtain

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum_{i=1}^{n} P(B \mid A_i)P(A_i)} \tag{1.45}$$

Note that the terms on the right-hand side are all conditioned on events A_i, while the term on the left is conditioned on B. Equation (*1.45*) is sometimes referred to as *Bayes' theorem*.

1.8 INDEPENDENT EVENTS

Two events A and B are said to be (*statistically*) *independent* if and only if

$$P(A \cap B) = P(A)P(B) \tag{1.46}$$

It follows immediately that if A and B are independent, then by Eqs. (*1.39*) and (*1.40*),

$$P(A \mid B) = P(A) \qquad \text{and} \qquad P(B \mid A) = P(B) \tag{1.47}$$

If two events A and B are independent, then it can be shown that A and $\bar{B}$ are also independent; that is (Prob. 1.53),

$$P(A \cap \bar{B}) = P(A)P(\bar{B}) \tag{1.48}$$

Then

$$P(A \mid \bar{B}) = \frac{P(A \cap \bar{B})}{P(\bar{B})} = P(A) \tag{1.49}$$

Thus, if A is independent of B, then the probability of A's occurrence is unchanged by information as to whether or not B has occurred. Three events A, B, C are said to be independent if and only if

$$\begin{aligned} P(A \cap B \cap C) &= P(A)P(B)P(C) \\ P(A \cap B) &= P(A)P(B) \\ P(A \cap C) &= P(A)P(C) \\ P(B \cap C) &= P(B)P(C) \end{aligned} \tag{1.50}$$

We may also extend the definition of independence to more than three events. The events $A_1, A_2, \ldots, A_n$ are independent if and only if for every subset $\{A_{i_1}, A_{i_2}, \ldots, A_{i_k}\}$ $(2 \le k \le n)$ of these events,

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k}) \tag{1.51}$$

Finally, we define an infinite set of events to be independent if and only if every finite subset of these events is independent.

To distinguish between the mutual exclusiveness (or disjointness) and independence of a collection of events we summarize as follows:

1. If $\{A_i, i = 1, 2, \ldots, n\}$ is a sequence of mutually exclusive events, then

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) \tag{1.52}$$

2. If $\{A_i, i = 1, 2, \ldots, n\}$ is a sequence of independent events, then

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i) \tag{1.53}$$

and a similar equality holds for any subcollection of the events.

Solved Problems

SAMPLE SPACE AND EVENTS

1.1. Consider a random experiment of tossing a coin three times.

(*a*) Find the sample space S_1 if we wish to observe the exact sequences of heads and tails obtained.

(*b*) Find the sample space S_2 if we wish to observe the number of heads in the three tosses.

(*a*) The sampling space S_1 is given by

$$S_1 = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

where, for example, HTH indicates a head on the first and third throws and a tail on the second throw. There are eight sample points in S_1.

(*b*) The sampling space S_2 is given by

$$S_2 = \{0, 1, 2, 3\}$$

where, for example, the outcome 2 indicates that two heads were obtained in the three tosses. The sample space S_2 contains four sample points.

1.2. Consider an experiment of drawing two cards at random from a bag containing four cards marked with the integers 1 through 4.

(*a*) Find the sample space S_1 of the experiment if the first card is replaced before the second is drawn.

(*b*) Find the sample space S_2 of the experiment if the first card is not replaced.

(*a*) The sample space S_1 contains 16 ordered pairs (i, j), $1 \le i \le 4$, $1 \le j \le 4$, where the first number indicates the first number drawn. Thus,

$$S_1 = \left\{\begin{matrix} (1, 1) & (1, 2) & (1, 3) & (1, 4) \\ (2, 1) & (2, 2) & (2, 3) & (2, 4) \\ (3, 1) & (3, 2) & (3, 3) & (3, 4) \\ (4, 1) & (4, 2) & (4, 3) & (4, 4) \end{matrix}\right\}$$