Introduction to Probability Theory

1.1 Introduction

Any realistic model of a real-world phenomenon must take into account the possibility of randomness. That is, more often than not, the quantities we are interested in will not be predictable in advance but, rather, will exhibit an inherent variation that should be taken into account by the model. This is usually accomplished by allowing the model to be probabilistic in nature. Such a model is, naturally enough, referred to as a probability model.

The majority of the chapters of this book will be concerned with different probability models of natural phenomena. Clearly, in order to master both the "model building" and the subsequent analysis of these models, we must have a certain knowledge of basic probability theory. The remainder of this chapter, as well as the next two chapters, will be concerned with a study of this subject.

1.2 Sample Space and Events

Suppose that we are about to perform an experiment whose outcome is not predictable in advance. However, while the outcome of the experiment will not be known in advance, let us suppose that the set of all possible outcomes is known. This set of all possible outcomes of an experiment is known as the *sample space* of the experiment and is denoted by S.

Introduction to Probability Models, ISBN: 9780123756862

Some examples are the following.

1. If the experiment consists of the flipping of a coin, then

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

where H means that the outcome of the toss is a head and T that it is a tail.

2. If the experiment consists of rolling a die, then the sample space is

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

where the outcome i means that i appeared on the die, $i = 1, 2, 3, 4, 5, 6$.

3. If the experiments consists of flipping two coins, then the sample space consists of the following four points:

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

The outcome will be (H, H) if both coins come up heads; it will be (H, T) if the first coin comes up heads and the second comes up tails; it will be (T, H) if the first comes up tails and the second heads; and it will be (T, T) if both coins come up tails.

4. If the experiment consists of rolling two dice, then the sample space consists of the following 36 points:

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

where the outcome (i, j) is said to occur if i appears on the first die and j on the second die.

5. If the experiment consists of measuring the lifetime of a car, then the sample space consists of all nonnegative real numbers. That is,

$$S = [0, \infty)^*$$ ■

Any subset E of the sample space S is known as an *event*. Some examples of events are the following.

1′. In Example (1), if $E = \{H\}$, then E is the event that a head appears on the flip of the coin. Similarly, if $E = \{T\}$, then E would be the event that a tail appears.

2′. In Example (2), if $E = \{1\}$, then E is the event that one appears on the roll of the die. If $E = \{2, 4, 6\}$, then E would be the event that an even number appears on the roll.

* The set (a, b) is defined to consist of all points x such that $a < x < b$. The set $[a, b]$ is defined to consist of all points x such that $a \leqslant x \leqslant b$. The sets $(a, b]$ and $[a, b)$ are defined, respectively, to consist of all points x such that $a < x \leqslant b$ and all points x such that $a \leqslant x < b$.

3′. In Example (3), if $E = \{(H, H),\ (H, T)\}$, then E is the event that a head appears on the first coin.
4′. In Example (4), if $E = \{(1, 6),\ (2, 5),\ (3, 4),\ (4, 3),\ (5, 2),\ (6, 1)\}$, then E is the event that the sum of the dice equals seven.
5′. In Example (5), if $E = (2, 6)$, then E is the event that the car lasts between two and six years. ■

We say that the event E occurs when the outcome of the experiment lies in E. For any two events E and F of a sample space S we define the new event $E \cup F$ to consist of all outcomes that are either in E or in F or in both E and F. That is, the event $E \cup F$ will occur if *either* E or F occurs. For example, in (1) if $E = \{H\}$ and $F = \{T\}$, then

$$E \cup F = \{H,\ T\}$$

That is, $E \cup F$ would be the whole sample space S. In (2) if $E = \{1, 3, 5\}$ and $F = \{1, 2, 3\}$, then

$$E \cup F = \{1, 2, 3, 5\}$$

and thus $E \cup F$ would occur if the outcome of the die is 1 or 2 or 3 or 5. The event $E \cup F$ is often referred to as the *union* of the event E and the event F.

For any two events E and F, we may also define the new event EF, sometimes written $E \cap F$, and referred to as the *intersection* of E and F, as follows. EF consists of all outcomes which are *both* in E and in F. That is, the event EF will occur only if both E and F occur. For example, in (2) if $E = \{1, 3, 5\}$ and $F = \{1, 2, 3\}$, then

$$EF = \{1, 3\}$$

and thus EF would occur if the outcome of the die is either 1 or 3. In Example (1) if $E = \{H\}$ and $F = \{T\}$, then the event EF would not consist of any outcomes and hence could not occur. To give such an event a name, we shall refer to it as the null event and denote it by Ø. (That is, Ø refers to the event consisting of no outcomes.) If $EF = \text{Ø}$, then E and F are said to be *mutually exclusive*.

We also define unions and intersections of more than two events in a similar manner. If $E_1, E_2, \ldots$ are events, then the union of these events, denoted by $\bigcup_{n=1}^{\infty} E_n$, is defined to be the event that consists of all outcomes that are in E_n for at least one value of $n = 1, 2, \ldots$. Similarly, the intersection of the events E_n, denoted by $\bigcap_{n=1}^{\infty} E_n$, is defined to be the event consisting of those outcomes that are in all of the events $E_n, n = 1, 2, \ldots$.

Finally, for any event E we define the new event E^c, referred to as the *complement* of E, to consist of all outcomes in the sample space S that are not in E. That is, E^c will occur if and only if E does not occur. In Example (4) if $E = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$, then E^c will occur if the sum of the dice does not equal seven. Also note that since the experiment must result in some outcome, it follows that $S^c = \emptyset$.

1.3 Probabilities Defined on Events

Consider an experiment whose sample space is S. For each event E of the sample space S, we assume that a number $P(E)$ is defined and satisfies the following three conditions:

(i) $0 \leqslant P(E) \leqslant 1$.
(ii) $P(S) = 1$.
(iii) For any sequence of events $E_1, E_2, \ldots$ that are mutually exclusive, that is, events for which $E_n E_m = \emptyset$ when $n \neq m$, then

$$P\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} P(E_n)$$

We refer to $P(E)$ as the probability of the event E.

Example 1.1 In the coin tossing example, if we assume that a head is equally likely to appear as a tail, then we would have

$$P(\{H\}) = P(\{T\}) = \tfrac{1}{2}$$

On the other hand, if we had a biased coin and felt that a head was twice as likely to appear as a tail, then we would have

$$P(\{H\}) = \tfrac{2}{3}, \qquad P(\{T\}) = \tfrac{1}{3}$$ ■

Example 1.2 In the die tossing example, if we supposed that all six numbers were equally likely to appear, then we would have

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = \tfrac{1}{6}$$

From (iii) it would follow that the probability of getting an even number would equal

$$\begin{aligned} P(\{2,4,6\}) &= P(\{2\}) + P(\{4\}) + P(\{6\}) \\ &= \tfrac{1}{2} \end{aligned}$$ ■

Remark We have chosen to give a rather formal definition of probabilities as being functions defined on the events of a sample space. However, it turns out that these probabilities have a nice intuitive property. Namely, if our experiment is repeated over and over again then (with probability 1) the proportion of time that event E occurs will just be $P(E)$.

Since the events E and E^c are always mutually exclusive and since $E \cup E^c = S$ we have by (ii) and (iii) that

$$1 = P(S) = P(E \cup E^c) = P(E) + P(E^c)$$

or

$$P(E^c) = 1 - P(E) \tag{1.1}$$

In words, Equation (1.1) states that the probability that an event does not occur is one minus the probability that it does occur.

We shall now derive a formula for $P(E \cup F)$, the probability of all outcomes either in E or in F. To do so, consider $P(E) + P(F)$, which is the probability of all outcomes in E plus the probability of all points in F. Since any outcome that is in both E and F will be counted twice in $P(E) + P(F)$ and only once in $P(E \cup F)$, we must have

$$P(E) + P(F) = P(E \cup F) + P(EF)$$

or equivalently

$$P(E \cup F) = P(E) + P(F) - P(EF) \tag{1.2}$$

Note that when E and F are mutually exclusive (that is, when $EF = \varnothing$), then Equation (1.2) states that

$$\begin{aligned} P(E \cup F) &= P(E) + P(F) - P(\varnothing) \\ &= P(E) + P(F) \end{aligned}$$

a result which also follows from condition (iii). (Why is $P(\varnothing) = 0$?)

Example 1.3 Suppose that we toss two coins, and suppose that we assume that each of the four outcomes in the sample space

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

is equally likely and hence has probability $\frac{1}{4}$. Let

$$E = \{(H, H), (H, T)\} \quad \text{and} \quad F = \{(H, H), (T, H)\}$$

That is, E is the event that the first coin falls heads, and F is the event that the second coin falls heads.

By Equation (1.2) we have that $P(E \cup F)$, the probability that either the first or the second coin falls heads, is given by

$$\begin{aligned}P(E \cup F) &= P(E) + P(F) - P(EF)\\ &= \tfrac{1}{2} + \tfrac{1}{2} - P(\{H, H\})\\ &= 1 - \tfrac{1}{4} = \tfrac{3}{4}\end{aligned}$$

This probability could, of course, have been computed directly since

$$P(E \cup F) = P(\{H, H), (H, T), (T, H)\}) = \tfrac{3}{4} \qquad \blacksquare$$

We may also calculate the probability that any one of the three events E or F or G occurs. This is done as follows:

$$P(E \cup F \cup G) = P((E \cup F) \cup G)$$

which by Equation (1.2) equals

$$P(E \cup F) + P(G) - P((E \cup F)G)$$

Now we leave it for you to show that the events $(E \cup F)G$ and $EG \cup FG$ are equivalent, and hence the preceding equals

$$\begin{aligned}&P(E \cup F \cup G)\\ &\quad= P(E) + P(F) - P(EF) + P(G) - P(EG \cup FG)\\ &\quad= P(E) + P(F) - P(EF) + P(G) - P(EG) - P(FG) + P(EGFG)\\ &\quad= P(E) + P(F) + P(G) - P(EF) - P(EG) - P(FG) + P(EFG) \qquad (1.3)\end{aligned}$$

In fact, it can be shown by induction that, for any n events $E_1, E_2, E_3, \dots, E_n$,

$$\begin{aligned}P(E_1 \cup E_2 \cup \cdots \cup E_n) = \sum_i P(E_i) - \sum_{i<j} P(E_iE_j) + \sum_{i<j<k} P(E_iE_jE_k)\\ - \sum_{i<j<k<l} P(E_iE_jE_kE_l)\\ + \cdots + (-1)^{n+1} P(E_1E_2 \cdots E_n) \qquad (1.4)\end{aligned}$$

In words, Equation (1.4), known as the *inclusion-exclusion identity*, states that the probability of the union of n events equals the sum of the probabilities of these events taken one at a time minus the sum of the probabilities of these events taken two at a time plus the sum of the probabilities of these events taken three at a time, and so on.

1.4 Conditional Probabilities

Suppose that we toss two dice and that each of the 36 possible outcomes is equally likely to occur and hence has probability $\frac{1}{36}$. Suppose that we observe that the first die is a four. Then, given this information, what is the probability that the sum of the two dice equals six? To calculate this probability we reason as follows: Given that the initial die is a four, it follows that there can be at most six possible outcomes of our experiment, namely, $(4, 1)$, $(4, 2)$, $(4, 3)$, $(4, 4)$, $(4, 5)$, and $(4, 6)$. Since each of these outcomes originally had the same probability of occurring, they should still have equal probabilities. That is, given that the first die is a four, then the (conditional) probability of each of the outcomes $(4, 1)$, $(4, 2)$, $(4, 3)$, $(4, 4)$, $(4, 5)$, $(4, 6)$ is $\frac{1}{6}$ while the (conditional) probability of the other 30 points in the sample space is 0. Hence, the desired probability will be $\frac{1}{6}$.

If we let E and F denote, respectively, the event that the sum of the dice is six and the event that the first die is a four, then the probability just obtained is called the conditional probability that E occurs given that F has occurred and is denoted by

$$P(E|F)$$

A general formula for $P(E|F)$ that is valid for all events E and F is derived in the same manner as the preceding. Namely, if the event F occurs, then in order for E to occur it is necessary for the actual occurrence to be a point in both E and in F, that is, it must be in EF. Now, because we know that F has occurred, it follows that F becomes our new sample space and hence the probability that the event EF occurs will equal the probability of EF relative to the probability of F. That is,

$$P(E|F) = \frac{P(EF)}{P(F)} \tag{1.5}$$

Note that Equation (1.5) is only well defined when $P(F) > 0$ and hence $P(E|F)$ is only defined when $P(F) > 0$.

Example 1.4 Suppose cards numbered one through ten are placed in a hat, mixed up, and then one of the cards is drawn. If we are told that the number on the drawn card is at least five, then what is the conditional probability that it is ten?

Solution: Let E denote the event that the number of the drawn card is ten, and let F be the event that it is at least five. The desired probability is $P(E|F)$. Now, from Equation (1.5)

$$P(E|F) = \frac{P(EF)}{P(F)}$$

However, $EF = E$ since the number of the card will be both ten and at least five if and only if it is number ten. Hence,

$$P(E|F) = \frac{\frac{1}{10}}{\frac{6}{10}} = \frac{1}{6} \qquad \blacksquare$$

Example 1.5 A family has two children. What is the conditional probability that both are boys given that at least one of them is a boy? Assume that the sample space S is given by $S = \{(b,b),(b,g),(g,b),(g,g)\}$, and all outcomes are equally likely. ((b,g) means, for instance, that the older child is a boy and the younger child a girl.)

Solution: Letting B denote the event that both children are boys, and A the event that at least one of them is a boy, then the desired probability is given by

$$\begin{aligned} P(B|A) &= \frac{P(BA)}{P(A)} \\ &= \frac{P(\{(b,b)\})}{P(\{(b,b),(b,g),(g,b)\})} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3} \qquad \blacksquare \end{aligned}$$

Example 1.6 Bev can either take a course in computers or in chemistry. If Bev takes the computer course, then she will receive an A grade with probability $\frac{1}{2}$; if she takes the chemistry course then she will receive an A grade with probability $\frac{1}{3}$. Bev decides to base her decision on the flip of a fair coin. What is the probability that Bev will get an A in chemistry?

Solution: If we let C be the event that Bev takes chemistry and A denote the event that she receives an A in whatever course she takes, then the desired probability is $P(AC)$. This is calculated by using Equation (1.5) as follows:

$$\begin{aligned} P(AC) &= P(C)P(A|C) \\ &= \tfrac{1}{2}\tfrac{1}{3} = \tfrac{1}{6} \qquad \blacksquare \end{aligned}$$

Example 1.7 Suppose an urn contains seven black balls and five white balls. We draw two balls from the urn without replacement. Assuming that each ball in the urn is equally likely to be drawn, what is the probability that both drawn balls are black?

Solution: Let F and E denote, respectively, the events that the first and second balls drawn are black. Now, given that the first ball selected is black, there are six remaining black balls and five white balls, and so $P(E|F) = \frac{6}{11}$. As $P(F)$ is clearly $\frac{7}{12}$, our desired probability is

$$\begin{aligned} P(EF) &= P(F)P(E|F) \\ &= \tfrac{7}{12}\tfrac{6}{11} = \tfrac{42}{132} \qquad \blacksquare \end{aligned}$$

Example 1.8 Suppose that each of three men at a party throws his hat into the center of the room. The hats are first mixed up and then each man randomly selects a hat. What is the probability that none of the three men selects his own hat?

Solution: We shall solve this by first calculating the complementary probability that at least one man selects his own hat. Let us denote by E_i, $i = 1, 2, 3$, the event that the ith man selects his own hat. To calculate the probability $P(E_1 \cup E_2 \cup E_3)$, we first note that

$$\begin{aligned} P(E_i) &= \tfrac{1}{3}, \qquad i = 1, 2, 3 \\ P(E_i E_j) &= \tfrac{1}{6}, \qquad i \neq j \\ P(E_1 E_2 E_3) &= \tfrac{1}{6} \end{aligned} \tag{1.6}$$

To see why Equation (1.6) is correct, consider first

$$P(E_i E_j) = P(E_i) P(E_j | E_i)$$

Now $P(E_i)$, the probability that the ith man selects his own hat, is clearly $\frac{1}{3}$ since he is equally likely to select any of the three hats. On the other hand, given that the ith man has selected his own hat, then there remain two hats that the jth man may select, and as one of these two is his own hat, it follows that with probability $\frac{1}{2}$ he will select it. That is, $P(E_j | E_i) = \frac{1}{2}$ and so

$$P(E_i E_j) = P(E_i) P(E_j | E_i) = \tfrac{1}{3} \tfrac{1}{2} = \tfrac{1}{6}$$

To calculate $P(E_1 E_2 E_3)$ we write

$$\begin{aligned} P(E_1 E_2 E_3) &= P(E_1 E_2) P(E_3 | E_1 E_2) \\ &= \tfrac{1}{6} P(E_3 | E_1 E_2) \end{aligned}$$

However, given that the first two men get their own hats it follows that the third man must also get his own hat (since there are no other hats left). That is, $P(E_3 | E_1 E_2) = 1$ and so

$$P(E_1 E_2 E_3) = \tfrac{1}{6}$$

Now, from Equation (1.4) we have that

$$\begin{aligned} P(E_1 \cup E_2 \cup E_3) &= P(E_1) + P(E_2) + P(E_3) - P(E_1 E_2) \\ &\quad - P(E_1 E_3) - P(E_2 E_3) + P(E_1 E_2 E_3) \\ &= 1 - \tfrac{1}{2} + \tfrac{1}{6} \\ &= \tfrac{2}{3} \end{aligned}$$

Hence, the probability that none of the men selects his own hat is $1-\frac{2}{3}=\frac{1}{3}$. ■

1.5 Independent Events

Two events E and F are said to be *independent* if

$$P(EF) = P(E)P(F)$$

By Equation (1.5) this implies that E and F are independent if

$$P(E|F) = P(E)$$

(which also implies that $P(F|E) = P(F)$). That is, E and F are independent if knowledge that F has occurred does not affect the probability that E occurs. That is, the occurrence of E is independent of whether or not F occurs.

Two events E and F that are not independent are said to be *dependent*.

Example 1.9 Suppose we toss two fair dice. Let E_1 denote the event that the sum of the dice is six and F denote the event that the first die equals four. Then

$$P(E_1F) = P(\{4, 2\}) = \frac{1}{36}$$

while

$$P(E_1)P(F) = \frac{5}{36}\frac{1}{6} = \frac{5}{216}$$

and hence E_1 and F are not independent. Intuitively, the reason for this is clear for if we are interested in the possibility of throwing a six (with two dice), then we will be quite happy if the first die lands four (or any of the numbers 1, 2, 3, 4, 5) because then we still have a possibility of getting a total of six. On the other hand, if the first die landed six, then we would be unhappy as we would no longer have a chance of getting a total of six. In other words, our chance of getting a total of six depends on the outcome of the first die and hence E_1 and F cannot be independent.

Let E_2 be the event that the sum of the dice equals seven. Is E_2 independent of F? The answer is yes since

$$P(E_2F) = P(\{(4, 3)\}) = \frac{1}{36}$$

while

$$P(E_2)P(F) = \frac{1}{6}\frac{1}{6} = \frac{1}{36}$$

We leave it for you to present the intuitive argument why the event that the sum of the dice equals seven is independent of the outcome on the first die. ■

The definition of independence can be extended to more than two events. The events $E_1, E_2, \ldots, E_n$ are said to be independent if for every subset $E_{1'}, E_{2'}, \ldots, E_{r'}$, $r \leqslant n$, of these events

$$P(E_{1'}E_{2'} \cdots E_{r'}) = P(E_{1'})P(E_{2'}) \cdots P(E_{r'})$$

Intuitively, the events $E_1, E_2, \ldots, E_n$ are independent if knowledge of the occurrence of any of these events has no effect on the probability of any other event.

Example 1.10 (Pairwise Independent Events That Are Not Independent) Let a ball be drawn from an urn containing four balls, numbered 1, 2, 3, 4. Let $E = \{1, 2\}$, $F = \{1, 3\}$, $G = \{1, 4\}$. If all four outcomes are assumed equally likely, then

$$\begin{aligned} P(EF) &= P(E)P(F) = \tfrac{1}{4}, \\ P(EG) &= P(E)P(G) = \tfrac{1}{4}, \\ P(FG) &= P(F)P(G) = \tfrac{1}{4} \end{aligned}$$

However,

$$\tfrac{1}{4} = P(EFG) \neq P(E)P(F)P(G)$$

Hence, even though the events E, F, G are pairwise independent, they are not jointly independent. ■

Example 1.11 There are r players, with player i initially having n_i units, $n_i > 0, i = 1, \ldots, r$. At each stage, two of the players are chosen to play a game, with the winner of the game receiving 1 unit from the loser. Any player whose fortune drops to 0 is eliminated, and this continues until a single player has all $n \equiv \sum_{i=1}^{r} n_i$ units, with that player designated as the victor. Assuming that the results of successive games are independent, and that each game is equally likely to be won by either of its two players, find the probability that player i is the victor.

Solution: To begin, suppose that there are n players, with each player initially having 1 unit. Consider player i. Each stage she plays will be equally likely to result in her either winning or losing 1 unit, with the results from each stage being independent. In addition, she will continue to play stages until her fortune becomes either 0 or n. Because this is the same for all players, it follows that each player has the same chance of being the victor. Consequently, each player has player probability $1/n$ of being the victor. Now, suppose these n players are divided into r teams, with team i containing n_i players, $i = 1, \ldots, r$. That is, suppose players $1, \ldots, n_1$ constitute team 1, players $n_1 + 1, \ldots, n_1 + n_2$ constitute team 2 and so on. Then the probability that the victor is a member of team i is n_i/n. But because team i initially has a total fortune of n_i units, $i = 1, \ldots, r$, and each game played by members of different teams results in the fortune of the winner's team increasing by 1 and that of the loser's team decreasing by 1, it is easy to see that the probability that the victor is from team i is exactly the desired probability. Moreover, our argument also shows that the result is true no matter how the choices of the players in each stage are made. ■

Suppose that a sequence of experiments, each of which results in either a "success" or a "failure," is to be performed. Let $E_i, i \geqslant 1$, denote the event that the ith experiment results in a success. If, for all $i_1, i_2, \ldots, i_n$,

$$P(E_{i_1} E_{i_2} \cdots E_{i_n}) = \prod_{j=1}^{n} P(E_{i_j})$$

we say that the sequence of experiments consists of *independent trials*.

1.6 Bayes' Formula

Let E and F be events. We may express E as

$$E = EF \cup EF^c$$

because in order for a point to be in E, it must either be in both E and F, or it must be in E and not in F. Since EF and EF^c are mutually exclusive, we have that

$$\begin{aligned} P(E) &= P(EF) + P(EF^c) \\ &= P(E|F)P(F) + P(E|F^c)P(F^c) \\ &= P(E|F)P(F) + P(E|F^c)(1 - P(F)) \end{aligned} \tag{1.7}$$

Equation (1.7) states that the probability of the event E is a weighted average of the conditional probability of E given that F has occurred and the conditional probability of E given that F has not occurred, each conditional probability being given as much weight as the event on which it is conditioned has of occurring.

Example 1.12 Consider two urns. The first contains two white and seven black balls, and the second contains five white and six black balls. We flip a fair coin and then draw a ball from the first urn or the second urn depending on whether the outcome was heads or tails. What is the conditional probability that the outcome of the toss was heads given that a white ball was selected?

Solution: Let W be the event that a white ball is drawn, and let H be the event that the coin comes up heads. The desired probability $P(H|W)$ may be calculated as follows:

$$\begin{aligned} P(H|W) &= \frac{P(HW)}{P(W)} = \frac{P(W|H)P(H)}{P(W)} \\ &= \frac{P(W|H)P(H)}{P(W|H)P(H) + P(W|H^c)P(H^c)} \\ &= \frac{\frac{2}{9}\frac{1}{2}}{\frac{2}{9}\frac{1}{2} + \frac{5}{11}\frac{1}{2}} = \frac{22}{67} \end{aligned}$$

■

Example 1.13 In answering a question on a multiple-choice test a student either knows the answer or guesses. Let p be the probability that she knows the answer and $1 - p$ the probability that she guesses. Assume that a student who guesses at the answer will be correct with probability $1/m$, where m is the number of multiple-choice alternatives. What is the conditional probability that a student knew the answer to a question given that she answered it correctly?

Solution: Let C and K denote respectively the event that the student answers the question correctly and the event that she actually knows the answer. Now

$$\begin{aligned} P(K|C) &= \frac{P(KC)}{P(C)} = \frac{P(C|K)P(K)}{P(C|K)P(K) + P(C|K^c)P(K^c)} \\ &= \frac{p}{p + (1/m)(1-p)} \\ &= \frac{mp}{1 + (m-1)p} \end{aligned}$$

Thus, for example, if $m = 5, p = \frac{1}{2}$, then the probability that a student knew the answer to a question she correctly answered is $\frac{5}{6}$. ■

Example 1.14 A laboratory blood test is 95 percent effective in detecting a certain disease when it is, in fact, present. However, the test also yields a "false positive" result for 1 percent of the healthy persons tested. (That is, if a healthy person is tested, then, with probability 0.01, the test result will imply he has the disease.) If 0.5 percent of the population actually has the disease, what is the probability a person has the disease given that his test result is positive?

Solution: Let D be the event that the tested person has the disease, and E the event that his test result is positive. The desired probability $P(D|E)$ is obtained by

$$\begin{aligned} P(D|E) &= \frac{P(DE)}{P(E)} = \frac{P(E|D)P(D)}{P(E|D)P(D) + P(E|D^c)P(D^c)} \\ &= \frac{(0.95)(0.005)}{(0.95)(0.005) + (0.01)(0.995)} \\ &= \frac{95}{294} \approx 0.323 \end{aligned}$$

Thus, only 32 percent of those persons whose test results are positive actually have the disease. ■

Equation (1.7) may be generalized in the following manner. Suppose that $F_1, F_2, \ldots, F_n$ are mutually exclusive events such that $\bigcup_{i=1}^n F_i = S$. In other words, exactly one of the events $F_1, F_2, \ldots, F_n$ will occur. By writing

$$E = \bigcup_{i=1}^{n} EF_i$$

and using the fact that the events EF_i, $i = 1, \ldots, n$, are mutually exclusive, we obtain that

$$\begin{aligned} P(E) &= \sum_{i=1}^{n} P(EF_i) \\ &= \sum_{i=1}^{n} P(E|F_i)P(F_i) \end{aligned} \tag{1.8}$$

Thus, Equation (1.8) shows how, for given events $F_1, F_2, \ldots, F_n$ of which one and only one must occur, we can compute $P(E)$ by first "conditioning" upon which one of the F_i occurs. That is, it states that $P(E)$ is equal to a weighted average of $P(E|F_i)$, each term being weighted by the probability of the event on which it is conditioned.

Suppose now that E has occurred and we are interested in determining which one of the F_j also occurred. By Equation (1.8) we have that

$$\begin{aligned} P(F_j|E) &= \frac{P(EF_j)}{P(E)} \\ &= \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^{n} P(E|F_i)P(F_i)} \end{aligned} \tag{1.9}$$

Equation (1.9) is known as *Bayes' formula*.

Example 1.15 You know that a certain letter is equally likely to be in any one of three different folders. Let α_i be the probability that you will find your letter upon making a quick examination of folder i if the letter is, in fact, in folder i, $i = 1, 2, 3$. (We may have $\alpha_i < 1$.) Suppose you look in folder 1 and do not find the letter. What is the probability that the letter is in folder 1?

Solution: Let F_i, $i = 1, 2, 3$ be the event that the letter is in folder i; and let E be the event that a search of folder 1 does not come up with the letter. We desire $P(F_1|E)$. From Bayes' formula we obtain

$$\begin{aligned} P(F_1|E) &= \frac{P(E|F_1)P(F_1)}{\sum_{i=1}^{3} P(E|F_i)P(F_i)} \\ &= \frac{(1-\alpha_1)\frac{1}{3}}{(1-\alpha_1)\frac{1}{3} + \frac{1}{3} + \frac{1}{3}} = \frac{1-\alpha_1}{3-\alpha_1} \end{aligned}$$

■

Exercises

1. A box contains three marbles: one red, one green, and one blue. Consider an experiment that consists of taking one marble from the box then replacing it in the box and drawing a second marble from the box. What is the sample space? If, at all times, each marble in the box is equally likely to be selected, what is the probability of each point in the sample space?

*2. Repeat Exercise 1 when the second marble is drawn without replacing the first marble.

3. A coin is to be tossed until a head appears twice in a row. What is the sample space for this experiment? If the coin is fair, what is the probability that it will be tossed exactly four times?

4. Let E, F, G be three events. Find expressions for the events that of E, F, G
   (a) only F occurs,
   (b) both E and F but not G occur,
   (c) at least one event occurs,
   (d) at least two events occur,
   (e) all three events occur,
   (f) none occurs,
   (g) at most one occurs,
   (h) at most two occur.

*5. An individual uses the following gambling system at Las Vegas. He bets \$1 that the roulette wheel will come up red. If he wins, he quits. If he loses then he makes the same bet a second time only this time he bets \$2; and then regardless of the outcome, quits. Assuming that he has a probability of $\frac{1}{2}$ of winning each bet, what is the probability that he goes home a winner? Why is this system not used by everyone?

6. Show that $E(F \cup G) = EF \cup EG$.

7. Show that $(E \cup F)^c = E^cF^c$.

8. If $P(E) = 0.9$ and $P(F) = 0.8$, show that $P(EF) \geqslant 0.7$. In general, show that

$$P(EF) \geqslant P(E) + P(F) - 1$$

This is known as Bonferroni's inequality.

*9. We say that $E \subset F$ if every point in E is also in F. Show that if $E \subset F$, then

$$P(F) = P(E) + P(FE^c) \geqslant P(E)$$

10. Show that

$$P\left(\bigcup_{i=1}^{n} E_i\right) \leqslant \sum_{i=1}^{n} P(E_i)$$

This is known as Boole's inequality.

Hint: Either use Equation (1.2) and mathematical induction, or else show that $\bigcup_{i=1}^{n} E_i = \bigcup_{i=1}^{n} F_i$, where $F_1 = E_1$, $F_i = E_i \bigcap_{j=1}^{i-1} E_j^c$, and use property (iii) of a probability.

11. If two fair dice are tossed, what is the probability that the sum is i, $i = 2, 3, \ldots, 12$?

12. Let E and F be mutually exclusive events in the sample space of an experiment. Suppose that the experiment is repeated until either event E or event F occurs. What does the sample space of this new super experiment look like? Show that the probability that event E occurs before event F is $P(E)/\,[P(E) + P(F)]$.

Hint: Argue that the probability that the original experiment is performed n times and E appears on the nth time is $P(E) \times (1-p)^{n-1}$, $n = 1, 2, \ldots$, where $p = P(E) + P(F)$. Add these probabilities to get the desired answer.

13. The dice game craps is played as follows. The player throws two dice, and if the sum is seven or eleven, then she wins. If the sum is two, three, or twelve, then she loses. If the sum is anything else, then she continues throwing until she either throws that number again (in which case she wins) or she throws a seven (in which case she loses). Calculate the probability that the player wins.

14. The probability of winning on a single toss of the dice is p. A starts, and if he fails, he passes the dice to B, who then attempts to win on her toss. They continue tossing the dice back and forth until one of them wins. What are their respective probabilities of winning?

15. Argue that $E = EF \cup EF^c$, $E \cup F = E \cup FE^c$.

16. Use Exercise 15 to show that $P(E \cup F) = P(E) + P(F) - P(EF)$.

*17. Suppose each of three persons tosses a coin. If the outcome of one of the tosses differs from the other outcomes, then the game ends. If not, then the persons start over and retoss their coins. Assuming fair coins, what is the probability that the game will end with the first round of tosses? If all three coins are biased and have probability $\frac{1}{4}$ of landing heads, what is the probability that the game will end at the first round?

18. Assume that each child who is born is equally likely to be a boy or a girl. If a family has two children, what is the probability that both are girls given that (a) the eldest is a girl, (b) at least one is a girl?

*19. Two dice are rolled. What is the probability that at least one is a six? If the two faces are different, what is the probability that at least one is a six?

20. Three dice are thrown. What is the probability the same number appears on exactly two of the three dice?

21. Suppose that 5 percent of men and 0.25 percent of women are color-blind. A color-blind person is chosen at random. What is the probability of this person being male? Assume that there are an equal number of males and females.

22. A and B play until one has 2 more points than the other. Assuming that each point is independently won by A with probability p, what is the probability they will play a total of $2n$ points? What is the probability that A will win?

23. For events $E_1, E_2, \ldots, E_n$ show that

$$P(E_1 E_2 \cdots E_n) = P(E_1)P(E_2|E_1)P(E_3|E_1 E_2) \cdots P(E_n|E_1 \cdots E_{n-1})$$

24. In an election, candidate A receives n votes and candidate B receives m votes, where $n > m$. Assume that in the count of the votes all possible orderings of the $n + m$ votes are equally likely. Let $P_{n,m}$ denote the probability that from the first vote on A is always in the lead. Find

(a) $P_{2,1}$ (b) $P_{3,1}$ (c) $P_{n,1}$ (d) $P_{3,2}$ (e) $P_{4,2}$
(f) $P_{n,2}$ (g) $P_{4,3}$ (h) $P_{5,3}$ (i) $P_{5,4}$
(j) Make a conjecture as to the value of $P_{n,m}$.

*25. Two cards are randomly selected from a deck of 52 playing cards.
(a) What is the probability they constitute a pair (that is, that they are of the same denomination)?
(b) What is the conditional probability they constitute a pair given that they are of different suits?

26. A deck of 52 playing cards, containing all 4 aces, is randomly divided into 4 piles of 13 cards each. Define events E_1, E_2, E_3, and E_4 as follows:

$E_1 = \{$the first pile has exactly 1 ace$\}$,
$E_2 = \{$the second pile has exactly 1 ace$\}$,
$E_3 = \{$the third pile has exactly 1 ace$\}$,
$E_4 = \{$the fourth pile has exactly 1 ace$\}$

Use Exercise 23 to find $P(E_1 E_2 E_3 E_4)$, the probability that each pile has an ace.

*27. Suppose in Exercise 26 we had defined the events E_i, $i = 1, 2, 3, 4$, by

$E_1 = \{$one of the piles contains the ace of spades$\}$,
$E_2 = \{$the ace of spades and the ace of hearts are in different piles$\}$,
$E_3 = \{$the ace of spades, the ace of hearts, and the ace of diamonds are in different piles$\}$,
$E_4 = \{$all 4 aces are in different piles$\}$

Now use Exercise 23 to find $P(E_1 E_2 E_3 E_4)$, the probability that each pile has an ace. Compare your answer with the one you obtained in Exercise 26.

28. If the occurrence of B makes A more likely, does the occurrence of A make B more likely?

29. Suppose that $P(E) = 0.6$. What can you say about $P(E|F)$ when
    (a) E and F are mutually exclusive?
    (b) $E \subset F$?
    (c) $F \subset E$?

*30. Bill and George go target shooting together. Both shoot at a target at the same time. Suppose Bill hits the target with probability 0.7, whereas George, independently, hits the target with probability 0.4.
    (a) Given that exactly one shot hit the target, what is the probability that it was George's shot?
    (b) Given that the target is hit, what is the probability that George hit it?

31. What is the conditional probability that the first die is six given that the sum of the dice is seven?

*32. Suppose all n men at a party throw their hats in the center of the room. Each man then randomly selects a hat. Show that the probability that none of the n men selects his own hat is

$$\frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - + \cdots \frac{(-1)^n}{n!}$$

Note that as $n \to \infty$ this converges to e^{-1}. Is this surprising?

33. In a class there are four freshman boys, six freshman girls, and six sophomore boys. How many sophomore girls must be present if sex and class are to be independent when a student is selected at random?

34. Mr. Jones has devised a gambling system for winning at roulette. When he bets, he bets on red, and places a bet only when the ten previous spins of the roulette have landed on a black number. He reasons that his chance of winning is quite large since the probability of eleven consecutive spins resulting in black is quite small. What do you think of this system?

35. A fair coin is continually flipped. What is the probability that the first four flips are
    (a) H, H, H, H?
    (b) T, H, H, H?
    (c) What is the probability that the pattern T, H, H, H occurs before the pattern H, H, H, H?

36. Consider two boxes, one containing one black and one white marble, the other, two black and one white marble. A box is selected at random and a marble is drawn at random from the selected box. What is the probability that the marble is black?

37. In Exercise 36, what is the probability that the first box was the one selected given that the marble is white?

38. Urn 1 contains two white balls and one black ball, while urn 2 contains one white ball and five black balls. One ball is drawn at random from urn 1 and placed in urn 2. A ball is then drawn from urn 2. It happens to be white. What is the probability that the transferred ball was white?

39. Stores A, B, and C have 50, 75, and 100 employees, and, respectively, 50, 60, and 70 percent of these are women. Resignations are equally likely among all employees,

regardless of sex. One employee resigns and this is a woman. What is the probability that she works in store C?

*40. (a) A gambler has in his pocket a fair coin and a two-headed coin. He selects one of the coins at random, and when he flips it, it shows heads. What is the probability that it is the fair coin?
(b) Suppose that he flips the same coin a second time and again it shows heads. Now what is the probability that it is the fair coin?
(c) Suppose that he flips the same coin a third time and it shows tails. Now what is the probability that it is the fair coin?

41. In a certain species of rats, black dominates over brown. Suppose that a black rat with two black parents has a brown sibling.
(a) What is the probability that this rat is a pure black rat (as opposed to being a hybrid with one black and one brown gene)?
(b) Suppose that when the black rat is mated with a brown rat, all five of their offspring are black. Now, what is the probability that the rat is a pure black rat?

42. There are three coins in a box. One is a two-headed coin, another is a fair coin, and the third is a biased coin that comes up heads 75 percent of the time. When one of the three coins is selected at random and flipped, it shows heads. What is the probability that it was the two-headed coin?

43. Suppose we have ten coins which are such that if the ith one is flipped then heads will appear with probability $i/10$, $i = 1, 2, \ldots, 10$. When one of the coins is randomly selected and flipped, it shows heads. What is the conditional probability that it was the fifth coin?

44. Urn 1 has five white and seven black balls. Urn 2 has three white and twelve black balls. We flip a fair coin. If the outcome is heads, then a ball from urn 1 is selected, while if the outcome is tails, then a ball from urn 2 is selected. Suppose that a white ball is selected. What is the probability that the coin landed tails?

*45. An urn contains b black balls and r red balls. One of the balls is drawn at random, but when it is put back in the urn c additional balls of the same color are put in with it. Now suppose that we draw another ball. Show that the probability that the first ball drawn was black given that the second ball drawn was red is $b/(b + r + c)$.

46. Three prisoners are informed by their jailer that one of them has been chosen at random to be executed, and the other two are to be freed. Prisoner A asks the jailer to tell him privately which of his fellow prisoners will be set free, claiming that there would be no harm in divulging this information, since he already knows that at least one will go free. The jailer refuses to answer this question, pointing out that if A knew which of his fellows were to be set free, then his own probability of being executed would rise from $\frac{1}{3}$ to $\frac{1}{2}$, since he would then be one of two prisoners. What do you think of the jailer's reasoning?

47. For a fixed event B, show that the collection $P(A|B)$, defined for all events A, satisfies the three conditions for a probability. Conclude from this that

$$P(A|B) = P(A|BC)P(C|B) + P(A|BC^c)P(C^c|B)$$

Then directly verify the preceding equation.

*48. Sixty percent of the families in a certain community own their own car, thirty percent own their own home, and twenty percent own both their own car and their own home. If a family is randomly chosen, what is the probability that this family owns a car or a house but not both?

References

Reference [2] provides a colorful introduction to some of the earliest developments in probability theory. References [3], [4], and [7] are all excellent introductory texts in modern probability theory. Reference [5] is the definitive work that established the axiomatic foundation of modern mathematical probability theory. Reference [6] is a nonmathematical introduction to probability theory and its applications, written by one of the greatest mathematicians of the eighteenth century.

[1] L. Breiman, "Probability," Addison-Wesley, Reading, Massachusetts, 1968.

[2] F. N. David, "Games, Gods, and Gambling," Hafner, New York, 1962.

[3] W. Feller, "An Introduction to Probability Theory and Its Applications," Vol. I, John Wiley, New York, 1957.

[4] B. V. Gnedenko, "Theory of Probability," Chelsea, New York, 1962.

[5] A. N. Kolmogorov, "Foundations of the Theory of Probability," Chelsea, New York, 1956.

[6] Marquis de Laplace, "A Philosophical Essay on Probabilities," 1825 (English Translation), Dover, New York, 1951.

[7] S. Ross, "A First Course in Probability," Eighth Edition, Prentice Hall, New Jersey, 2010.