

ELEMENTS OF PROBABILITY

3.1 INTRODUCTION

The concept of the probability of a particular event of an experiment is subject to various meanings or interpretations. For instance, if a geologist is quoted as saying that “there is a 60 percent chance of oil in a certain region,” we all probably have some intuitive idea as to what is being said. Indeed, most of us would probably interpret this statement in one of two possible ways: either by imagining that

1. the geologist feels that, over the long run, in 60 percent of the regions whose outward environmental conditions are very similar to the conditions that prevail in the region under consideration, there will be oil; or
2. the geologist believes that it is more likely that the region will contain oil than it is that it will not; and in fact .6 is a measure of the geologist’s belief in the hypothesis that the region will contain oil.

The two foregoing interpretations of the probability of an event are referred to as being the frequency interpretation and the subjective (or personal) interpretation of probability. In the *frequency interpretation*, the probability of a given outcome of an experiment is considered as being a “property” of that outcome. It is imagined that this property can be operationally determined by continual repetition of the experiment — the probability of the outcome will then be observable as being the proportion of the experiments that result in the outcome. This is the interpretation of probability that is most prevalent among scientists.

In the subjective interpretation, the probability of an outcome is not thought of as being a property of the outcome but rather is considered a statement about the beliefs of the person who is quoting the probability, concerning the chance that the outcome will occur. Thus, in this interpretation, probability becomes a subjective or personal concept and has no meaning outside of expressing one’s degree of belief. This interpretation of probability is often favored by philosophers and certain economic decision makers.

Regardless of which interpretation one gives to probability, however, there is a consensus that the mathematics of probability are the same in either case. For instance, if you think that the probability that it will rain tomorrow is .3 and you feel that the probability that it will be cloudy but without any rain is .2, then you should feel that the probability that it will either be cloudy or rainy is .5 independently of your individual interpretation of the concept of probability. In this chapter, we present the accepted rules, or axioms, used in probability theory. As a preliminary to this, however, we need to study the concept of the sample space and the events of an experiment.

3.2 SAMPLE SPACE AND EVENTS

Consider an experiment whose outcome is not predictable with certainty in advance. Although 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 . Some examples are the following.

1. If the outcome of an experiment consists in the determination of the sex of a newborn child, then

$$S = \{g, b\}$$

where the outcome g means that the child is a girl and b that it is a boy.

2. If the experiment consists of the running of a race among the seven horses having post positions 1, 2, 3, 4, 5, 6, 7, then

$$S = \{\text{all orderings of } (1, 2, 3, 4, 5, 6, 7)\}$$

The outcome (2, 3, 1, 6, 5, 4, 7) means, for instance, that the number 2 horse is first, then the number 3 horse, then the number 1 horse, and so on.

3. Suppose we are interested in determining the amount of dosage that must be given to a patient until that patient reacts positively. One possible sample space for this experiment is to let S consist of all the positive numbers. That is, let

$$S = (0, \infty)$$

where the outcome would be x if the patient reacts to a dosage of value x but not to any smaller dosage.

Any subset E of the sample space is known as an *event*. That is, an event is a set consisting of possible outcomes of the experiment. If the outcome of the experiment is contained in E , then we say that E has occurred. Some examples of events are the following.

In Example 1 if $E = \{g\}$, then E is the event that the child is a girl. Similarly, if $F = \{b\}$, then F is the event that the child is a boy.

In Example 2 if

$$E = \{\text{all outcomes in } S \text{ starting with a 3}\}$$

then E is the event that the number 3 horse wins the race.

For any two events E and F of a sample space S , we define the new event $E \cup F$, called the *union* of the events E and 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 instance, in Example 1 if $E = \{g\}$ and $F = \{b\}$, then $E \cup F = \{g, b\}$. That is, $E \cup F$ would be the whole sample space S . In Example 2 if $E = \{\text{all outcomes starting with 6}\}$ is the event that the number 6 horse wins and $F = \{\text{all outcomes having 6 in the second position}\}$ is the event that the number 6 horse comes in second, then $E \cup F$ is the event that the number 6 horse comes in either first or second.

Similarly, for any two events E and F , we may also define the new event EF , sometimes written as $E \cap F$, called the *intersection* of E and F , to consist of all outcomes that are in both E and F . That is, the event EF will occur only if both E and F occur. For instance, in Example 3 if $E = (0, 5)$ is the event that the required dosage is less than 5 and $F = (2, 10)$ is the event that it is between 2 and 10, then $EF = (2, 5)$ is the event that the required dosage is between 2 and 5. In Example 2 if $E = \{\text{all outcomes ending in 5}\}$ is the event that horse number 5 comes in last and $F = \{\text{all outcomes starting with 5}\}$ is the event that horse number 5 comes in first, then the event EF does not contain any outcomes and hence cannot occur. To give such an event a name, we shall refer to it as the null event and denote it by \emptyset . Thus \emptyset refers to the event consisting of no outcomes. If $EF = \emptyset$, implying that E and F cannot both occur, then E and F are said to be *mutually exclusive*.

For any event E , we define the 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 1 if $E = \{b\}$ is the event that the child is a boy, then $E^c = \{g\}$ is the event that it is a girl. Also note that since the experiment must result in some outcome, it follows that $S^c = \emptyset$.

For any two events E and F , if all of the outcomes in E are also in F , then we say that E is contained in F and write $E \subset F$ (or equivalently, $F \supset E$). Thus if $E \subset F$, then the occurrence of E necessarily implies the occurrence of F . If $E \subset F$ and $F \subset E$, then we say that E and F are equal (or identical) and we write $E = F$.

We can also define unions and intersections of more than two events. In particular, the union of the events E_1, E_2, \dots, E_n , denoted either by $E_1 \cup E_2 \cup \dots \cup E_n$ or by $\bigcup_1^n E_i$, is defined to be the event consisting of all outcomes that are in E_i for at least one $i = 1, 2, \dots, n$. Similarly, the intersection of the events $E_i, i = 1, 2, \dots, n$, denoted by $E_1 E_2 \cdots E_n$, is defined to be the event consisting of those outcomes that are in all of the events $E_i, i = 1, 2, \dots, n$. In other words, the union of the E_i occurs when *at least* one of the events E_i occurs; the intersection occurs when *all* of the events E_i occur.

3.3 VENN DIAGRAMS AND THE ALGEBRA OF EVENTS

A graphical representation of events that is very useful for illustrating logical relations among them is the *Venn diagram*. The sample space S is represented as consisting of all the points in a large rectangle, and the events E, F, G, \dots , are represented as consisting of all the points in given circles within the rectangle. Events of interest can then be *indicated* by shading appropriate regions of the diagram. For instance, in the three Venn diagrams shown in Figure 3.1, the shaded areas represent respectively the events $E \cup F$, EF , and E^c . The Venn diagram of Figure 3.2 indicates that $E \subset F$.

The operations of forming unions, intersections, and complements of events obey certain rules not dissimilar to the rules of algebra. We list a few of these.

Commutative law	$E \cup F = F \cup E$	$EF = FE$
Associative law	$(E \cup F) \cup G = E \cup (F \cup G)$	$(EF)G = E(FG)$
Distributive law	$(E \cup F)G = EG \cup FG$	$EF \cup G = (E \cup G)(F \cup G)$

These relations are verified by showing that any outcome that is contained in the event on the left side of the equality is also contained in the event on the right side and vice versa. One way of showing this is by means of Venn diagrams. For instance, the distributive law may be verified by the sequence of diagrams shown in Figure 3.3.

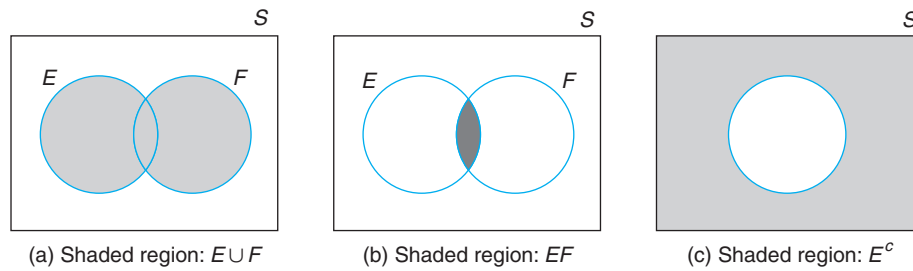


FIGURE 3.1 Venn diagrams.

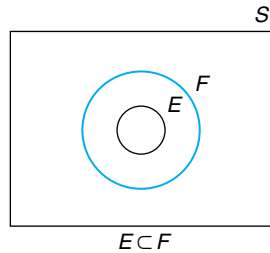


FIGURE 3.2 Venn diagram.

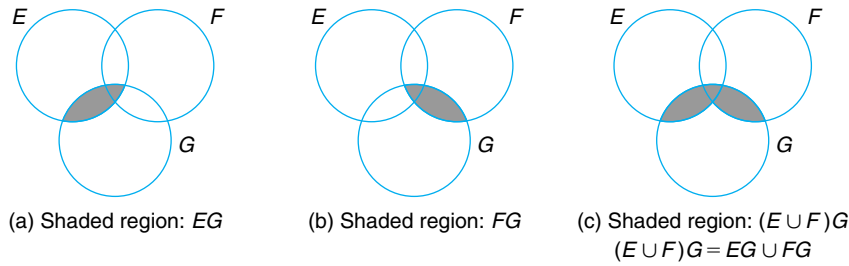


FIGURE 3.3 Proving the distributive law.

The following useful relationship between the three basic operations of forming unions, intersections, and complements of events is known as *DeMorgan's laws*.

$$(E \cup F)^c = E^c F^c$$

$$(EF)^c = E^c \cup F^c$$

3.4 AXIOMS OF PROBABILITY

It appears to be an empirical fact that if an experiment is continually repeated under the exact same conditions, then for any event E , the proportion of time that the outcome is contained in E approaches some constant value as the number of repetitions increases. For instance, if a coin is continually flipped, then the proportion of flips resulting in heads will approach some value as the number of flips increases. It is this constant limiting frequency that we often have in mind when we speak of the probability of an event.

From a purely mathematical viewpoint, we shall suppose that for each event E of an experiment having a sample space S there is a number, denoted by $P(E)$, that is in accord with the following three axioms:

AXIOM 1

$$0 \leq P(E) \leq 1$$

AXIOM 2

$$P(S) = 1$$

AXIOM 3

For any sequence of mutually exclusive events E_1, E_2, \dots (that is, events for which $E_i E_j = \emptyset$ when $i \neq j$),

$$P\left(\bigcup_{i=1}^n E_i\right) = \sum_{i=1}^n P(E_i), \quad n = 1, 2, \dots, \infty$$

We call $P(E)$ the probability of the event E .

Thus, Axiom 1 states that the probability that the outcome of the experiment is contained in E is some number between 0 and 1. Axiom 2 states that, with probability 1,

the outcome will be a member of the sample space S . Axiom 3 states that for any set of mutually exclusive events the probability that at least one of these events occurs is equal to the sum of their respective probabilities.

It should be noted that if we interpret $P(E)$ as the relative frequency of the event E when a large number of repetitions of the experiment are performed, then $P(E)$ would indeed satisfy the above axioms. For instance, the proportion (or frequency) of time that the outcome is in E is clearly between 0 and 1, and the proportion of time that it is in S is 1 (since all outcomes are in S). Also, if E and F have no outcomes in common, then the proportion of time that the outcome is in either E or F is the sum of their respective frequencies. As an illustration of this last statement, suppose the experiment consists of the rolling of a pair of dice and suppose that E is the event that the sum is 2, 3, or 12 and F is the event that the sum is 7 or 11. Then if outcome E occurs 11 percent of the time and outcome F 22 percent of the time, then 33 percent of the time the outcome will be either 2, 3, 12, 7, or 11.

These axioms will now be used to prove two simple propositions concerning probabilities. We first note that E and E^c are always mutually exclusive, and since $E \cup E^c = S$, we have by Axioms 2 and 3 that

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

Or equivalently, we have the following:

PROPOSITION 3.4.1

$$P(E^c) = 1 - P(E)$$

In other words, Proposition 3.4.1 states that the probability that an event does not occur is 1 minus the probability that it does occur. For instance, if the probability of obtaining a head on the toss of a coin is $\frac{3}{8}$, the probability of obtaining a tail must be $\frac{5}{8}$.

Our second proposition gives the relationship between the probability of the union of two events in terms of the individual probabilities and the probability of the intersection.

PROPOSITION 3.4.2

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

Proof

This proposition is most easily proven by the use of a Venn diagram as shown in Figure 3.4. As the regions I, II, and III are mutually exclusive, it follows that

$$P(E \cup F) = P(\text{I}) + P(\text{II}) + P(\text{III})$$

$$P(E) = P(\text{I}) + P(\text{II})$$

$$P(F) = P(\text{II}) + P(\text{III})$$

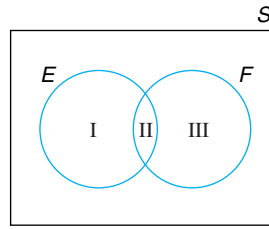


FIGURE 3.4 Venn diagram.

which shows that

$$P(E \cup F) = P(E) + P(F) - P(\text{II})$$

and the proof is complete since $\text{II} = EF$. ■

EXAMPLE 3.4a A total of 28 percent of males living in Nevada smoke cigarettes, 6 percent smoke cigars, and 3 percent smoke both cigars and cigarettes. What percentage of males smoke neither cigars nor cigarettes?

SOLUTION Let E be the event that a randomly chosen male is a cigarette smoker and let F be the event that he is a cigar smoker. Then, the probability this person is either a cigarette or a cigar smoker is

$$P(E \cup F) = P(E) + P(F) - P(EF) = .28 + .06 - .03 = .31$$

Thus the probability that the person is not a smoker is $1 - .31 = .69$, implying that 69 percent of males smoke neither cigarettes nor cigars. ■

The *odds* of an event A is defined by

$$\frac{P(A)}{P(A^c)} = \frac{P(A)}{1 - P(A)}$$

Thus the odds of an event A tells how much more likely it is that A occurs than that it does not occur. For instance, if $P(A) = 3/4$, then $P(A)/(1 - P(A)) = 3$, so the odds are 3. Consequently, it is 3 times as likely that A occurs as it is that it does not. (Common terminology is to say that the odds are 3 to 1 in favor of the event A .)

3.5 SAMPLE SPACES HAVING EQUALLY LIKELY OUTCOMES

For a large number of experiments, it is natural to assume that each point in the sample space is equally likely to occur. That is, for many experiments whose sample space S is a finite set, say $S = \{1, 2, \dots, N\}$, it is often natural to assume that

$$P(\{1\}) = P(\{2\}) = \dots = P(\{N\}) = p \quad (\text{say})$$

Now it follows from Axioms 2 and 3 that

$$1 = P(S) = P(\{1\}) + \cdots + P(\{N\}) = Np$$

which shows that

$$P(\{i\}) = p = 1/N$$

From this it follows from Axiom 3 that for any event E ,

$$P(E) = \frac{\text{Number of points in } E}{N}$$

In words, if we assume that each outcome of an experiment is equally likely to occur, then the probability of any event E equals the proportion of points in the sample space that are contained in E .

Thus, to compute probabilities it is often necessary to be able to effectively count the number of different ways that a given event can occur. To do this, we will make use of the following rule.

BASIC PRINCIPLE OF COUNTING

Suppose that two experiments are to be performed. Then if experiment 1 can result in any one of m possible outcomes and if, for each outcome of experiment 1, there are n possible outcomes of experiment 2, then together there are mn possible outcomes of the two experiments.

Proof of the Basic Principle

The basic principle can be proven by enumerating all the possible outcomes of the two experiments as follows:

$$\begin{aligned} &(1, 1), (1, 2), \dots, (1, n) \\ &(2, 1), (2, 2), \dots, (2, n) \\ &\vdots \\ &(m, 1), (m, 2), \dots, (m, n) \end{aligned}$$

where we say that the outcome is (i, j) if experiment 1 results in its i th possible outcome and experiment 2 then results in the j th of its possible outcomes. Hence, the set of possible outcomes consists of m rows, each row containing n elements, which proves the result. ■

EXAMPLE 3.5a Two balls are “randomly drawn” from a bowl containing 6 white and 5 black balls. What is the probability that one of the drawn balls is white and the other black?

SOLUTION If we regard the order in which the balls are selected as being significant, then as the first drawn ball may be any of the 11 and the second any of the remaining 10, it follows that the sample space consists of $11 \cdot 10 = 110$ points. Furthermore, there are $6 \cdot 5 = 30$ ways in which the first ball selected is white and the second black, and similarly there are $5 \cdot 6 = 30$ ways in which the first ball is black and the second white. Hence, assuming that “randomly drawn” means that each of the 110 points in the sample space is equally likely to occur, then we see that the desired probability is

$$\frac{30 + 30}{110} = \frac{6}{11} \quad \blacksquare$$

When there are more than two experiments to be performed the basic principle can be generalized as follows:

Generalized Basic Principle of Counting

If r experiments that are to be performed are such that the first one may result in any of n_1 possible outcomes, and if for each of these n_1 possible outcomes there are n_2 possible outcomes of the second experiment, and if for each of the possible outcomes of the first two experiments there are n_3 possible outcomes of the third experiment, and if, \dots , then there are a total of $n_1 \cdot n_2 \cdots n_r$ possible outcomes of the r experiments.

As an illustration of this, let us determine the number of different ways n distinct objects can be arranged in a linear order. For instance, how many different ordered arrangements of the letters a, b, c are possible? By direct enumeration we see that there are 6; namely, $abc, acb, bac, bca, cab, cba$. Each one of these ordered arrangements is known as a *permutation*. Thus, there are 6 possible permutations of a set of 3 objects. This result could also have been obtained from the basic principle, since the first object in the permutation can be any of the 3, the second object in the permutation can then be chosen from any of the remaining 2, and the third object in the permutation is then chosen from the remaining one. Thus, there are $3 \cdot 2 \cdot 1 = 6$ possible permutations.

Suppose now that we have n objects. Similar reasoning shows that there are

$$n(n-1)(n-2) \cdots 3 \cdot 2 \cdot 1$$

different permutations of the n objects. It is convenient to introduce the notation $n!$, which is read “ n factorial,” for the foregoing expression. That is,

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

Thus, for instance, $1! = 1$, $2! = 2 \cdot 1 = 2$, $3! = 3 \cdot 2 \cdot 1 = 6$, $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$, and so on. It is convenient to define $0! = 1$.

EXAMPLE 3.5b Mr. Jones has 10 books that he is going to put on his bookshelf. Of these, 4 are mathematics books, 3 are chemistry books, 2 are history books, and 1 is a language book. Jones wants to arrange his books so that all the books dealing with the same subject are together on the shelf. How many different arrangements are possible?

SOLUTION There are $4! 3! 2! 1!$ arrangements such that the mathematics books are first in line, then the chemistry books, then the history books, and then the language book. Similarly, for each possible ordering of the subjects, there are $4! 3! 2! 1!$ possible arrangements. Hence, as there are $4!$ possible orderings of the subjects, the desired answer is $4! 4! 3! 2! 1! = 6,912$. ■

EXAMPLE 3.5c A class in probability theory consists of 6 men and 4 women. An exam is given and the students are ranked according to their performance. Assuming that no two students obtain the same score, (a) how many different rankings are possible? (b) If all rankings are considered equally likely, what is the probability that women receive the top 4 scores?

SOLUTION

- (a) Because each ranking corresponds to a particular ordered arrangement of the 10 people, we see the answer to this part is $10! = 3,628,800$.
- (b) Because there are $4!$ possible rankings of the women among themselves and $6!$ possible rankings of the men among themselves, it follows from the basic principle that there are $(6!)(4!) = (720)(24) = 17,280$ possible rankings in which the women receive the top 4 scores. Hence, the desired probability is

$$\frac{6!4!}{10!} = \frac{4 \cdot 3 \cdot 2 \cdot 1}{10 \cdot 9 \cdot 8 \cdot 7} = \frac{1}{210} \quad \blacksquare$$

Suppose now that we are interested in determining the number of different groups of r objects that could be formed from a total of n objects. For instance, how many different groups of three could be selected from the five items A, B, C, D, E ? To answer this, reason as follows. Since there are 5 ways to select the initial item, 4 ways to then select the next item, and 3 ways to then select the final item, there are thus $5 \cdot 4 \cdot 3$ ways of selecting the group of 3 when the order in which the items are selected is relevant. However, since every group of 3, say the group consisting of items A, B , and C , will be counted 6 times (that is, all of the permutations $ABC, ACB, BAC, BCA, CAB, CBA$ will be counted when the order of selection is relevant), it follows that the total number of different groups that can be formed is $(5 \cdot 4 \cdot 3)/(3 \cdot 2 \cdot 1) = 10$.

In general, as $n(n-1) \cdots (n-r+1)$ represents the number of different ways that a group of r items could be selected from n items when the order of selection is considered

relevant (since the first one selected can be any one of the n , and the second selected any one of the remaining $n - 1$, etc.), and since each group of r items will be counted $r!$ times in this count, it follows that the number of different groups of r items that could be formed from a set of n items is

$$\frac{n(n-1) \cdots (n-r+1)}{r!} = \frac{n!}{(n-r)!r!}$$

NOTATION AND TERMINOLOGY

We define $\binom{n}{r}$, for $r \leq n$, by

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

and call $\binom{n}{r}$ the number of *combinations* of n objects taken r at a time.

Thus $\binom{n}{r}$ represents the number of different groups of size r that can be selected from a set of size n when the order of selection is not considered relevant. For example, there are

$$\binom{8}{2} = \frac{8 \cdot 7}{2 \cdot 1} = 28$$

different groups of size 2 that can be chosen from a set of 8 people, and

$$\binom{10}{2} = \frac{10 \cdot 9}{2 \cdot 1} = 45$$

different groups of size 2 that can be chosen from a set of 10 people. Also, since $0! = 1$, note that

$$\binom{n}{0} = \binom{n}{n} = 1$$

EXAMPLE 3.5d A committee of size 5 is to be selected from a group of 6 men and 9 women. If the selection is made randomly, what is the probability that the committee consists of 3 men and 2 women?

SOLUTION Let us assume that “randomly selected” means that each of the $\binom{15}{5}$ possible combinations is equally likely to be selected. Hence, since there are $\binom{6}{3}$ possible choices of 3 men and $\binom{9}{2}$ possible choices of 2 women, it follows from the basic principle of counting that the desired probability is given by

$$\frac{\binom{6}{3} \binom{9}{2}}{\binom{15}{5}} = \frac{240}{1001} \quad \blacksquare$$

EXAMPLE 3.5e From a set of n items a random sample of size k is to be selected. What is the probability a given item will be among the k selected?

SOLUTION Because there is $\binom{1}{1}$ way of choosing the given item and $\binom{n-1}{k-1}$ different choices of $k-1$ of the other $n-1$ items, it follows from the basic principle of counting that there are $\binom{1}{1}\binom{n-1}{k-1} = \binom{n-1}{k-1}$ different subsets of k of the n items that include the given item. As there are a total of $\binom{n}{k}$ different choices of k of the n items, it follows that the probability that a particular item is among the k selected is

$$\binom{n-1}{k-1} / \binom{n}{k} = \frac{(n-1)!}{(n-k)!(k-1)!} / \frac{n!}{(n-k)!k!} = \frac{k}{n} \quad \blacksquare$$

EXAMPLE 3.5f A basketball team consists of 6 black and 6 white players. The players are to be paired in groups of two for the purpose of determining roommates. If the pairings are done at random, what is the probability that none of the black players will have a white roommate?

SOLUTION Let us start by imagining that the 6 pairs are numbered — that is, there is a first pair, a second pair, and so on. Since there are $\binom{12}{2}$ different choices of a first pair; and for each choice of a first pair there are $\binom{10}{2}$ different choices of a second pair; and for each choice of the first 2 pairs there are $\binom{8}{2}$ choices for a third pair; and so on, it follows from the generalized basic principle of counting that there are

$$\binom{12}{2}\binom{10}{2}\binom{8}{2}\binom{6}{2}\binom{4}{2}\binom{2}{2} = \frac{12!}{(2!)^6}$$

ways of dividing the players into a *first* pair, a *second* pair, and so on. Hence there are $(12)!/2^6 6!$ ways of dividing the players into 6 (unordered) pairs of 2 each. Furthermore, since there are, by the same reasoning, $6!/2^3 3!$ ways of pairing the white players among themselves and $6!/2^3 3!$ ways of pairing the black players among themselves, it follows that there are $(6!/2^3 3!)^2$ pairings that do not result in any black–white roommate pairs. Hence, if the pairings are done at random (so that all outcomes are equally likely), then the desired probability is

$$\left(\frac{6!}{2^3 3!}\right)^2 / \frac{(12)!}{2^6 6!} = \frac{5}{231} = .0216$$

Hence, there are roughly only two chances in a hundred that a random pairing will not result in any of the white and black players rooming together. \blacksquare

EXAMPLE 3.5g If n people are present in a room, what is the probability that no two of them celebrate their birthday on the same day of the year? How large need n be so that this probability is less than $\frac{1}{2}$?

SOLUTION Because each person can celebrate his or her birthday on any one of 365 days, there are a total of $(365)^n$ possible outcomes. (We are ignoring the possibility of someone

having been born on February 29.) Furthermore, there are $(365)(364)(363) \cdots (365 - n + 1)$ possible outcomes that result in no two of the people having the same birthday. This is so because the first person could have any one of 365 birthdays, the next person any of the remaining 364 days, the next any of the remaining 363, and so on. Hence, assuming that each outcome is equally likely, we see that the desired probability is

$$\frac{(365)(364)(363) \cdots (365 - n + 1)}{(365)^n}$$

It is a rather surprising fact that when $n \geq 23$, this probability is less than $\frac{1}{2}$. That is, if there are 23 or more people in a room, then the probability that at least two of them have the same birthday exceeds $\frac{1}{2}$. Many people are initially surprised by this result, since 23 seems so small in relation to 365, the number of days of the year. However, every pair of individuals has probability $\frac{365}{(365)^2} = \frac{1}{365}$ of having the same birthday, and in a group of 23 people there are $\binom{23}{2} = 253$ different pairs of individuals. Looked at this way, the result no longer seems so surprising. ■

3.6 CONDITIONAL PROBABILITY

In this section, we introduce one of the most important concepts in all of probability theory — that of conditional probability. Its importance is twofold. In the first place, we are often interested in calculating probabilities when some partial information concerning the result of the experiment is available, or in recalculating them in light of additional information. In such situations, the desired probabilities are conditional ones. Second, as a kind of a bonus, it often turns out that the easiest way to compute the probability of an event is to first “condition” on the occurrence or nonoccurrence of a secondary event.

As an illustration of a conditional probability, suppose that one rolls a pair of dice. The sample space S of this experiment can be taken to be the following set of 36 outcomes

$$S = \{(i, j), \quad i = 1, 2, 3, 4, 5, 6, \quad j = 1, 2, 3, 4, 5, 6\}$$

where we say that the outcome is (i, j) if the first die lands on side i and the second on side j . Suppose now that each of the 36 possible outcomes is equally likely to occur and thus has probability $\frac{1}{36}$. (In such a situation we say that the dice are fair.) Suppose further that we observe that the first die lands on side 3. Then, given this information, what is the probability that the sum of the two dice equals 8? To calculate this probability, we reason as follows: Given that the initial die is a 3, there can be at most 6 possible outcomes of our experiment, namely, $(3, 1)$, $(3, 2)$, $(3, 3)$, $(3, 4)$, $(3, 5)$, and $(3, 6)$. In addition, because 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 3, then the (conditional) probability of each of the outcomes $(3, 1)$, $(3, 2)$, $(3, 3)$, $(3, 4)$, $(3, 5)$, $(3, 6)$ is $\frac{1}{6}$, whereas the (conditional) probability of the other 30 points in the sample space is 0. Hence, the desired probability will be $\frac{1}{6}$.

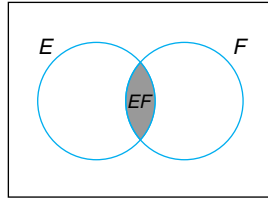


FIGURE 3.5 $P(E|F) = \frac{P(EF)}{P(F)}$.

If we let E and F denote, respectively, the event that the sum of the dice is 8 and the event that the first die is a 3, then the probability just obtained is called the conditional probability of E 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 just described. Namely, if the event F occurs, then in order for E to occur it is necessary that the actual occurrence be a point in both E and F ; that is, it must be in EF . However, because we know that F has occurred, it follows that we can regard F as the 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,

$$P(E|F) = \frac{P(EF)}{P(F)} \quad (3.6.1)$$

Note that Equation 3.6.1 is well defined only when $P(F) > 0$ and hence $P(E|F)$ is defined only when $P(F) > 0$. (See Figure 3.5.)

The definition of conditional probability given by Equation 3.6.1 is consistent with the interpretation of probability as being a long-run relative frequency. To see this, suppose that a large number n of repetitions of the experiment are performed. Then, since $P(F)$ is the long-run proportion of experiments in which F occurs, it follows that F will occur approximately $nP(F)$ times. Similarly, in approximately $nP(EF)$ of these experiments, both E and F will occur. Hence, of the approximately $nP(F)$ experiments whose outcome is in F , approximately $nP(EF)$ of them will also have their outcome in E . That is, for those experiments whose outcome is in F , the proportion whose outcome is also in E is approximately

$$\frac{nP(EF)}{nP(F)} = \frac{P(EF)}{P(F)}$$

Since this approximation becomes exact as n becomes larger and larger, it follows that (3.6.1) gives the appropriate definition of the conditional probability of E given that F has occurred.

EXAMPLE 3.6a A bin contains 5 defective (that immediately fail when put in use), 10 partially defective (that fail after a couple of hours of use), and 25 acceptable transistors.

A transistor is chosen at random from the bin and put into use. If it does not immediately fail, what is the probability it is acceptable?

SOLUTION Since the transistor did not immediately fail, we know that it is not one of the 5 defectives and so the desired probability is:

$$\begin{aligned} P\{\text{acceptable}|\text{not defective}\} &= \frac{P\{\text{acceptable, not defective}\}}{P\{\text{not defective}\}} \\ &= \frac{P\{\text{acceptable}\}}{P\{\text{not defective}\}} \end{aligned}$$

where the last equality follows since the transistor will be both acceptable and not defective if it is acceptable. Hence, assuming that each of the 40 transistors is equally likely to be chosen, we obtain that

$$P\{\text{acceptable}|\text{not defective}\} = \frac{25/40}{35/40} = 5/7$$

It should be noted that we could also have derived this probability by working directly with the reduced sample space. That is, since we know that the chosen transistor is not defective, the problem reduces to computing the probability that a transistor, chosen at random from a bin containing 25 acceptable and 10 partially defective transistors, is acceptable. This is clearly equal to $\frac{25}{35}$. ■

EXAMPLE 3.6b The organization that Jones works for is running a father–son dinner for those employees having at least one son. Each of these employees is invited to attend along with his youngest son. If Jones is known to have two children, what is the conditional probability that they are both boys given that he is invited to the dinner? 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 younger child is a boy and the older child is a girl].

SOLUTION The knowledge that Jones has been invited to the dinner is equivalent to knowing that he has at least one son. Hence, letting B denote the event that both children are boys, and A the event that at least one of them is a boy, we have that the desired probability $P(B|A)$ is given by

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

Many readers incorrectly reason that the conditional probability of two boys given at least one is $\frac{1}{2}$, as opposed to the correct $\frac{1}{3}$, since they reason that the Jones child not attending the dinner is equally likely to be a boy or a girl. Their mistake, however, is in assuming that these two possibilities are equally likely. Remember that initially there were four equally likely outcomes. Now the information that at least one child is a boy is equivalent to knowing that the outcome is not (g, g) . Hence we are left with the three equally likely outcomes (b, b) , (b, g) , (g, b) , thus showing that the Jones child not attending the dinner is twice as likely to be a girl as a boy. ■

By multiplying both sides of Equation 3.6.1 by $P(F)$ we obtain that

$$P(EF) = P(F)P(E|F) \quad (3.6.2)$$

In words, Equation 3.6.2 states that the probability that both E and F occur is equal to the probability that F occurs multiplied by the conditional probability of E given that F occurred. Equation 3.6.2 is often quite useful in computing the probability of the intersection of events. This is illustrated by the following example.

EXAMPLE 3.6c Ms. Perez figures that there is a 30 percent chance that her company will set up a branch office in Phoenix. If it does, she is 60 percent certain that she will be made manager of this new operation. What is the probability that Perez will be a Phoenix branch office manager?

SOLUTION If we let B denote the event that the company sets up a branch office in Phoenix and M the event that Perez is made the Phoenix manager, then the desired probability is $P(BM)$, which is obtained as follows:

$$\begin{aligned} P(BM) &= P(B)P(M|B) \\ &= (.3)(.6) \\ &= .18 \end{aligned}$$

Hence, there is an 18 percent chance that Perez will be the Phoenix manager. ■

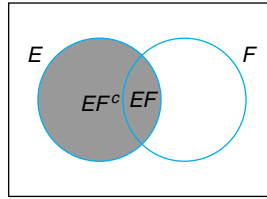
3.7 BAYES' FORMULA

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

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

for, in order for a point to be in E , it must either be in both E and F or be in E but not in F . (See Figure 3.6.) As EF and EF^c are clearly mutually exclusive, we have by Axiom 3 that

$$\begin{aligned} P(E) &= P(EF) + P(EF^c) \\ &= P(E|F)P(F) + P(E|F^c)P(F^c) \\ &= P(E|F)P(F) + P(E|F^c)[1 - P(F)] \end{aligned} \quad (3.7.1)$$

FIGURE 3.6 $E = EF \cup EF^c$.

Equation 3.7.1 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, with each conditional probability being given as much weight as the event it is conditioned on has of occurring. It is an extremely useful formula, for its use often enables us to determine the probability of an event by first “conditioning” on whether or not some second event has occurred. That is, there are many instances where it is difficult to compute the probability of an event directly, but it is straightforward to compute it once we know whether or not some second event has occurred.

EXAMPLE 3.7a An insurance company believes that people can be divided into two classes — those that are accident prone and those that are not. Their statistics show that an accident-prone person will have an accident at some time within a fixed 1-year period with probability .4, whereas this probability decreases to .2 for a non-accident-prone person. If we assume that 30 percent of the population is accident prone, what is the probability that a new policy holder will have an accident within a year of purchasing a policy?

SOLUTION We obtain the desired probability by first conditioning on whether or not the policy holder is accident prone. Let A_1 denote the event that the policy holder will have an accident within a year of purchase; and let A denote the event that the policy holder is accident prone. Hence, the desired probability, $P(A_1)$, is given by

$$\begin{aligned} P(A_1) &= P(A_1|A)P(A) + P(A_1|A^c)P(A^c) \\ &= (.4)(.3) + (.2)(.7) = .26 \quad \blacksquare \end{aligned}$$

In the next series of examples, we will indicate how to reevaluate an initial probability assessment in the light of additional (or new) information. That is, we will show how to incorporate new information with an initial probability assessment to obtain an updated probability.

EXAMPLE 3.7b Twins can either be identical or fraternal. Identical, also called monozygotic, twins form when a single fertilized egg splits into two genetically identical parts. Consequently, identical twins always have the same set of genes. Fraternal, also called dizygotic, twins develop when two separate eggs are fertilized and implant in the uterus. The genetic connection of fraternal twins is no more or less the same as siblings born at separate times. A Los Angeles county scientist wishing to know the current fraction of

twin pairs born in the county that are identical twins has assigned a county statistician to study this issue. The statistician initially requested each hospital in the county to record all twin births, indicating whether the resulting twins were identical or not. The hospitals, however, told her that to determine whether newborn twins were identical was not a simple task, as it involved the permission of the twins's parents to perform complicated and expensive DNA studies that the hospitals could not afford. After some deliberation, the statistician just asked the hospitals for data listing all twin births along with an indication as to whether the twins were of the same sex. When such data indicated that approximately 64 percent of twin births were same-sexed, the statistician declared that approximately 28 percent of all twins were identical. How did she come to this conclusion?

SOLUTION The statistician reasoned that identical twins are always of the same sex, whereas fraternal twins, having the same relationship to each other as any pair of siblings, will have probability $\frac{1}{2}$ of being of the same sex. Letting I be the event that a pair of twins are identical, and SS be the event that a pair of twins are of the same sex, she computed the probability $P(SS)$ by conditioning on whether the twin pair was identical. This gave

$$P(SS) = P(SS|I)P(I) + P(SS|I^c)P(I^c)$$

or

$$P(SS) = 1 \times P(I) + \frac{1}{2} \times [1 - P(I)] = \frac{1}{2} + \frac{1}{2}P(I)$$

which, using that $P(SS) \approx .64$ yielded the result

$$P(I) \approx .28 \quad \blacksquare$$

EXAMPLE 3.7c Reconsider Example 3.7a and suppose that a new policy holder has an accident within a year of purchasing his policy. What is the probability that he is accident prone?

SOLUTION Initially, at the moment when the policy holder purchased his policy, we assumed there was a 30 percent chance that he was accident prone. That is, $P(A) = .3$. However, based on the fact that he has had an accident within a year, we now reevaluate his probability of being accident prone as follows.

$$\begin{aligned} P(A|A_1) &= \frac{P(AA_1)}{P(A_1)} \\ &= \frac{P(A)P(A_1|A)}{P(A_1)} \\ &= \frac{(.3)(.4)}{.26} = \frac{6}{13} = .4615 \quad \blacksquare \end{aligned}$$

EXAMPLE 3.7d In answering a question on a multiple-choice test, a student either knows the answer or she 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 events that the student answers the question correctly and the event that she actually knows the answer. To compute

$$P(K|C) = \frac{P(KC)}{P(C)}$$

we first note that

$$\begin{aligned} P(KC) &= P(K)P(C|K) \\ &= p \cdot 1 \\ &= p \end{aligned}$$

To compute the probability that the student answers correctly, we condition on whether or not she knows the answer. That is,

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

Hence, the desired probability is given by

$$P(K|C) = \frac{p}{p + (1/m)(1 - p)} = \frac{mp}{1 + (m - 1)p}$$

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 3.7e A laboratory blood test is 99 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 .01, the test result will imply he or she has the disease.) If .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

$$\begin{aligned}
 P(D|E) &= \frac{P(DE)}{P(E)} \\
 &= \frac{P(E|D)P(D)}{P(E|D)P(D) + P(E|D^c)P(D^c)} \\
 &= \frac{(.99)(.005)}{(.99)(.005) + (.01)(.995)} \\
 &= .3322
 \end{aligned}$$

Thus, only 33 percent of those persons whose test results are positive actually have the disease. Since many students are often surprised at this result (because they expected this figure to be much higher since the blood test seems to be a good one), it is probably worthwhile to present a second argument which, though less rigorous than the foregoing, is probably more revealing. We now do so.

Since .5 percent of the population actually has the disease, it follows that, on the average, 1 person out of every 200 tested will have it. The test will correctly confirm that this person has the disease with probability .99. Thus, on the average, out of every 200 persons tested, the test will correctly confirm that .99 person has the disease. On the other hand, out of the (on the average) 199 healthy people, the test will incorrectly state that (199) (.01) of these people have the disease. Hence, for every .99 diseased person that the test correctly states is ill, there are (on the average) 1.99 healthy persons that the test incorrectly states are ill. Hence, the proportion of time that the test result is correct when it states that a person is ill is

$$\frac{.99}{.99 + 1.99} = .3322 \quad \blacksquare$$

Equation 3.7.1 is also useful when one has to reassess one's (personal) probabilities in the light of additional information. For instance, consider the following examples.

EXAMPLE 3.7f At a certain stage of a criminal investigation, the inspector in charge is 60 percent convinced of the guilt of a certain suspect. Suppose now that a *new* piece of evidence that shows that the criminal has a certain characteristic (such as left-handedness, baldness, brown hair, etc.) is uncovered. If 20 percent of the population possesses this characteristic, how certain of the guilt of the suspect should the inspector now be if it turns out that the suspect is among this group?

SOLUTION Letting G denote the event that the suspect is guilty and C the event that he possesses the characteristic of the criminal, we have

$$P(G|C) = \frac{P(GC)}{P(C)}$$

Now

$$\begin{aligned} P(GC) &= P(G)P(C|G) \\ &= (.6)(1) \\ &= .6 \end{aligned}$$

To compute the probability that the suspect has the characteristic, we condition on whether or not he is guilty. That is,

$$\begin{aligned} P(C) &= P(C|G)P(G) + P(C|G^c)P(G^c) \\ &= (1)(.6) + (.2)(.4) \\ &= .68 \end{aligned}$$

where we have supposed that the probability of the suspect having the characteristic if he is, in fact, innocent is equal to .2, the proportion of the population possessing the characteristic. Hence

$$P(G|C) = \frac{60}{68} = .882$$

and so the inspector should now be 88 percent certain of the guilt of the suspect. ■

EXAMPLE 3.7f (continued) Let us now suppose that the new evidence is subject to different possible interpretations, and in fact only shows that it is 90 percent likely that the criminal possesses this certain characteristic. In this case, how likely would it be that the suspect is guilty (assuming, as before, that he has this characteristic)?

SOLUTION In this case, the situation is as before with the exception that the probability of the suspect having the characteristic given that he is guilty is now .9 (rather than 1). Hence,

$$\begin{aligned} P(G|C) &= \frac{P(GC)}{P(C)} \\ &= \frac{P(G)P(C|G)}{P(C|G)P(G) + P(C|G^c)P(G^c)} \\ &= \frac{(.6)(.9)}{(.9)(.6) + (.2)(.4)} \\ &= \frac{54}{62} = .871 \end{aligned}$$

which is slightly less than in the previous case (why?). ■

Equation 3.7.1 may be generalized in the following manner. Suppose that F_1, F_2, \dots, 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, \dots, F_n must occur. By writing

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

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

$$\begin{aligned} P(E) &= \sum_{i=1}^n P(EF_i) \\ &= \sum_{i=1}^n P(E|F_i)P(F_i) \end{aligned} \quad (3.7.2)$$

Thus, Equation 3.7.2 shows how, for given events F_1, F_2, \dots, F_n of which one and only one must occur, we can compute $P(E)$ by first “conditioning” on 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 F_j also occurred. By Equation 3.7.2, we have that

$$\begin{aligned} P(F_j|E) &= \frac{P(EF_j)}{P(E)} \\ &= \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^n P(E|F_i)P(F_i)} \end{aligned} \quad (3.7.3)$$

Equation 3.7.3 is known as *Bayes’ formula*, after the English philosopher Thomas Bayes. If we think of the events F_j as being possible “hypotheses” about some subject matter, then Bayes’ formula may be interpreted as showing us how opinions about these hypotheses held before the experiment [that is, the $P(F_j)$] should be modified by the evidence of the experiment.

EXAMPLE 3.7g A plane is missing and it is presumed that it was equally likely to have gone down in any of three possible regions. Let $1 - \alpha_i$ denote the probability the plane will be found upon a search of the i th region when the plane is, in fact, in that region, $i = 1, 2, 3$. (The constants α_i are called *overlook probabilities* because they represent the

probability of overlooking the plane; they are generally attributable to the geographical and environmental conditions of the regions.) What is the conditional probability that the plane is in the i th region, given that a search of region 1 is unsuccessful, $i = 1, 2, 3$?

SOLUTION Let R_i , $i = 1, 2, 3$, be the event that the plane is in region i ; and let E be the event that a search of region 1 is unsuccessful. From Bayes' formula, we obtain

$$\begin{aligned}
 P(R_1|E) &= \frac{P(ER_1)}{P(E)} \\
 &= \frac{P(E|R_1)P(R_1)}{\sum_{i=1}^3 P(E|R_i)P(R_i)} \\
 &= \frac{(\alpha_1)(1/3)}{(\alpha_1)(1/3) + (1)(1/3) + (1)(1/3)} \\
 &= \frac{\alpha_1}{\alpha_1 + 2}
 \end{aligned}$$

For $j = 2, 3$,

$$\begin{aligned}
 P(R_j|E) &= \frac{P(E|R_j)P(R_j)}{P(E)} \\
 &= \frac{(1)(1/3)}{(\alpha_1)1/3 + 1/3 + 1/3} \\
 &= \frac{1}{\alpha_1 + 2}, \quad j = 2, 3
 \end{aligned}$$

Thus, for instance, if $\alpha_1 = .4$, then the conditional probability that the plane is in region 1 given that a search of that region did not uncover it is $\frac{1}{6}$, whereas the conditional probabilities that it is in region 2 and that it is in region 3 are both equal to $\frac{1}{2.4} = \frac{5}{12}$. ■

3.8 INDEPENDENT EVENTS

The previous examples in this chapter show that $P(E|F)$, the conditional probability of E given F , is not generally equal to $P(E)$, the unconditional probability of E . In other words, knowing that F has occurred generally changes the chances of E 's occurrence. In the special cases where $P(E|F)$ does in fact equal $P(E)$, we say that E is independent of F . That is, E is independent of F if knowledge that F has occurred does not change the probability that E occurs.

Since $P(E|F) = P(EF)/P(F)$, we see that E is independent of F if

$$P(EF) = P(E)P(F) \tag{3.8.1}$$

Since this equation is symmetric in E and F , it shows that whenever E is independent of F so is F of E . We thus have the following.

Definition

Two events E and F are said to be *independent* if Equation 3.8.1 holds. Two events E and F that are not independent are said to be *dependent*.

EXAMPLE 3.8a A card is selected at random from an ordinary deck of 52 playing cards. If A is the event that the selected card is an ace and H is the event that it is a heart, then A and H are independent, since $P(AH) = \frac{1}{52}$, while $P(A) = \frac{4}{52}$ and $P(H) = \frac{13}{52}$. ■

EXAMPLE 3.8b If we let E denote the event that the next president is a Republican and F the event that there will be a major earthquake within the next year, then most people would probably be willing to assume that E and F are independent. However, there would probably be some controversy over whether it is reasonable to assume that E is independent of G , where G is the event that there will be a recession within the next two years. ■

We now show that if E is independent of F then E is also independent of F^c .

PROPOSITION 3.8.1 If E and F are independent, then so are E and F^c .

Proof

Assume that E and F are independent. Since $E = EF \cup EF^c$, and EF and EF^c are obviously mutually exclusive, we have that

$$\begin{aligned} P(E) &= P(EF) + P(EF^c) \\ &= P(E)P(F) + P(EF^c) \quad \text{by the independence of } E \text{ and } F \end{aligned}$$

or equivalently,

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

and the result is proven. ■

Thus if E is independent of F , then the probability of E 's occurrence is unchanged by information as to whether or not F has occurred.

Suppose now that E is independent of F and is also independent of G . Is E then necessarily independent of FG ? The answer, somewhat surprisingly, is no. Consider the following example.

EXAMPLE 3.8c Two fair dice are thrown. Let E_7 denote the event that the sum of the dice is 7. Let F denote the event that the first die equals 4 and let T be the event that the

second die equals 3. Now it can be shown (see Problem 36) that E_7 is independent of F and that E_7 is also independent of T ; but clearly E_7 is not independent of FT [since $P(E_7|FT) = 1]$. ■

It would appear to follow from the foregoing example that an appropriate definition of the independence of three events E , F , and G would have to go further than merely assuming that all of the $\binom{3}{2}$ pairs of events are independent. We are thus led to the following definition.

Definition

The three events E , F , and G are said to be independent if

$$\begin{aligned} P(EFG) &= P(E)P(F)P(G) \\ P(EF) &= P(E)P(F) \\ P(EG) &= P(E)P(G) \\ P(FG) &= P(F)P(G) \end{aligned}$$

It should be noted that if the events E , F , G are independent, then E will be independent of any event formed from F and G . For instance, E is independent of $F \cup G$ since

$$\begin{aligned} P(E(F \cup G)) &= P(EF \cup EG) \\ &= P(EF) + P(EG) - P(EFG) \\ &= P(E)P(F) + P(E)P(G) - P(E)P(FG) \\ &= P(E)[P(F) + P(G) - P(FG)] \\ &= P(E)P(F \cup G) \end{aligned}$$

Of course we may also extend the definition of independence to more than three events. The events E_1, E_2, \dots, E_n are said to be independent if for every subset $E_{1'}, E_{2'}, \dots, E_{r'}$, $r \leq n$, of these events

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

It is sometimes the case that the probability experiment under consideration consists of performing a sequence of subexperiments. For instance, if the experiment consists of continually tossing a coin, then we may think of each toss as being a subexperiment. In many cases it is reasonable to assume that the outcomes of any group of the subexperiments have no effect on the probabilities of the outcomes of the other subexperiments. If such is the case, then we say that the subexperiments are independent.

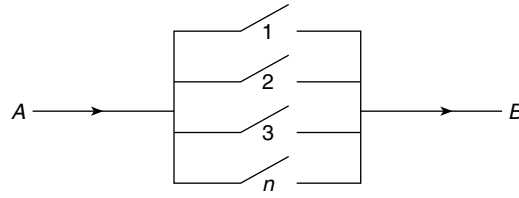


FIGURE 3.7 Parallel system: functions if current flows from A to B .

EXAMPLE 3.8d A system composed of n separate components is said to be a parallel system if it functions when at least one of the components functions. (See Figure 3.7.) For such a system, if component i , independent of other components, functions with probability p_i , $i = 1, \dots, n$, what is the probability the system functions?

SOLUTION Let A_i denote the event that component i functions. Then

$$\begin{aligned}
 P\{\text{system functions}\} &= 1 - P\{\text{system does not function}\} \\
 &= 1 - P\{\text{all components do not function}\} \\
 &= 1 - P(A_1^c A_2^c \cdots A_n^c) \\
 &= 1 - \prod_{i=1}^n (1 - p_i) \quad \text{by independence} \quad \blacksquare
 \end{aligned}$$

EXAMPLE 3.8e A set of k coupons, each of which is independently a type j coupon with probability p_j , $\sum_{j=1}^n p_j = 1$, is collected. Find the probability that the set contains a type j coupon given that it contains a type i , $i \neq j$.

SOLUTION Let A_r be the event that the set contains a type r coupon. Then

$$P(A_j|A_i) = \frac{P(A_j A_i)}{P(A_i)}$$

To compute $P(A_i)$ and $P(A_j A_i)$, consider the probability of their complements:

$$\begin{aligned}
 P(A_i) &= 1 - P(A_i^c) \\
 &= 1 - P\{\text{no coupon is type } i\} \\
 &= 1 - (1 - p_i)^k \\
 P(A_i A_j) &= 1 - P(A_i^c \cup A_j^c) \\
 &= 1 - [P(A_i^c) + P(A_j^c) - P(A_i^c A_j^c)] \\
 &= 1 - (1 - p_i)^k - (1 - p_j)^k + P\{\text{no coupon is type } i \text{ or type } j\} \\
 &= 1 - (1 - p_i)^k - (1 - p_j)^k + (1 - p_i - p_j)^k
 \end{aligned}$$

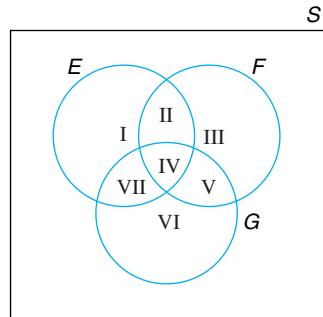
where the final equality follows because each of the k coupons is, independently, neither of type i or of type j with probability $1 - p_i - p_j$. Consequently,

$$P(A_j|A_i) = \frac{1 - (1 - p_i)^k - (1 - p_j)^k + (1 - p_i - p_j)^k}{1 - (1 - p_i)^k} \quad \blacksquare$$

Problems

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. Describe the sample space. Repeat for the case in which the second marble is drawn without first replacing the first marble.
2. An experiment consists of tossing a coin three times. What is the sample space of this experiment? Which event corresponds to the experiment resulting in more heads than tails?
3. Let $S = \{1, 2, 3, 4, 5, 6, 7\}$, $E = \{1, 3, 5, 7\}$, $F = \{7, 4, 6\}$, $G = \{1, 4\}$. Find
 - (a) EF ; (c) EG^c ; (e) $E^c(F \cup G)$;
 - (b) $E \cup FG$; (d) $EF^c \cup G$; (f) $EG \cup FG$.
4. Two dice are thrown. Let E be the event that the sum of the dice is odd, let F be the event that the first die lands on 1, and let G be the event that the sum is 5. Describe the events EF , $E \cup F$, FG , EF^c , EFG .
5. A system is composed of four components, each of which is either working or failed. Consider an experiment that consists of observing the status of each component, and let the outcome of the experiment be given by the vector (x_1, x_2, x_3, x_4) where x_i is equal to 1 if component i is working and is equal to 0 if component i is failed.
 - (a) How many outcomes are in the sample space of this experiment?
 - (b) Suppose that the system will work if components 1 and 2 are both working, or if components 3 and 4 are both working. Specify all the outcomes in the event that the system works.
 - (c) Let E be the event that components 1 and 3 are both failed. How many outcomes are contained in event E ?
6. Let E , F , G be three events. Find expressions for the events that of E , F , G
 - (a) only E occurs;
 - (b) both E and G but not F occur;
 - (c) at least one of the events occurs;

- (d) at least two of the events occur;
 (e) all three occur;
 (f) none of the events occurs;
 (g) at most one of them occurs;
 (h) at most two of them occur;
 (i) exactly two of them occur;
 (j) at most three of them occur.
7. Find simple expressions for the events
- (a) $E \cup E^c$;
 (b) EE^c ;
 (c) $(E \cup F)(E \cup F^c)$;
 (d) $(E \cup F)(E^c \cup F)(E \cup F^c)$;
 (e) $(E \cup F)(F \cup G)$.
8. Use Venn diagrams (or any other method) to show that
- (a) $EF \subset E, E \subset E \cup F$;
 (b) if $E \subset F$ then $F^c \subset E^c$;
 (c) the commutative laws are valid;
 (d) the associative laws are valid;
 (e) $F = FE \cup FE^c$;
 (f) $E \cup F = E \cup E^c F$;
 (g) DeMorgan's laws are valid.
9. For the following Venn diagram, describe in terms of E , F , and G the events denoted in the diagram by the Roman numerals I through VII.



10. Show that if $E \subset F$ then $P(E) \leq P(F)$. (Hint: Write F as the union of two mutually exclusive events, one of them being E .)
11. Prove Boole's inequality, namely that

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

12. If $P(E) = .9$ and $P(F) = .9$, show that $P(EF) \geq .8$. In general, prove Bonferroni's inequality, namely that

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

13. Prove that

(a) $P(EF^c) = P(E) - P(EF)$

(b) $P(E^cF^c) = 1 - P(E) - P(F) + P(EF)$

14. Show that the probability that exactly one of the events E or F occurs is equal to $P(E) + P(F) - 2P(EF)$.

15. Calculate $\binom{9}{3}$, $\binom{9}{6}$, $\binom{7}{2}$, $\binom{7}{5}$, $\binom{10}{7}$.

16. Show that

$$\binom{n}{r} = \binom{n}{n-r}$$

Now present a combinatorial argument for the foregoing by explaining why a choice of r items from a set of size n is equivalent to a choice of $n - r$ items from that set.

17. Show that

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$$

For a combinatorial argument, consider a set of n items and fix attention on one of these items. How many different sets of size r contain this item, and how many do not?

18. A group of 5 boys and 10 girls is lined up in random order — that is, each of the $15!$ permutations is assumed to be equally likely.
- What is the probability that the person in the 4th position is a boy?
 - What about the person in the 12th position?
 - What is the probability that a particular boy is in the 3rd position?
19. Consider a set of 23 unrelated people. Because each pair of people shares the same birthday with probability $1/365$, and there are $\binom{23}{2} = 253$ pairs, why isn't the probability that at least two people have the same birthday equal to $253/365$?
20. Suppose that distinct integer values are written on each of 3 cards. These cards are then randomly given the designations A, B, and C. The values on cards A and B are then compared. If the smaller of these values is then compared with the value on card C, what is the probability that it is also smaller than the value on card C?
21. There is a 60 percent chance that the event A will occur. If A does not occur, then there is a 10 percent chance that B will occur.

- (a) What is the probability that at least one of the events A or B occurs?
- (b) If A is the event that the democratic candidate wins the presidential election in 2012 and B is the event that there is a 6.2 or higher earthquake in Los Angeles sometime in 2013, what would you take as the probability that both A and B occur? What assumption are you making?
22. The sample mean of the annual salaries of a group of 100 accountants who work at a large accounting firm is \$130,000 with a sample standard deviation of \$20,000. If a member of this group is randomly chosen, what can we say about
- (a) the probability that his or her salary is between \$90,000 and \$170,000?
- (b) the probability that his or her salary exceeds \$150,000?
- Hint:* Use the Chebyshev inequality.
23. Of three cards, one is painted red on both sides; one is painted black on both sides; and one is painted red on one side and black on the other. A card is randomly chosen and placed on a table. If the side facing up is red, what is the probability that the other side is also red?
24. A couple has 2 children. What is the probability that both are girls if the eldest is a girl?
25. Fifty-two percent of the students at a certain college are females. Five percent of the students in this college are majoring in computer science. Two percent of the students are women majoring in computer science. If a student is selected at random, find the conditional probability that
- (a) this student is female, given that the student is majoring in computer science;
- (b) this student is majoring in computer science, given that the student is female.
26. A total of 500 married working couples were polled about their annual salaries, with the following information resulting.

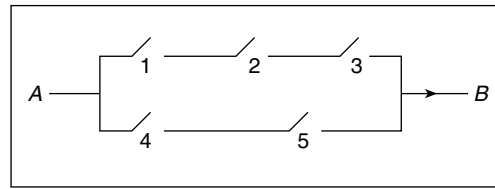
Wife	Husband	
	Less than \$50,000	More than \$50,000
Less than \$50,000	212	198
More than \$50,000	36	54

Thus, for instance, in 36 of the couples the wife earned more and the husband earned less than \$50,000. If one of the couples is randomly chosen, what is

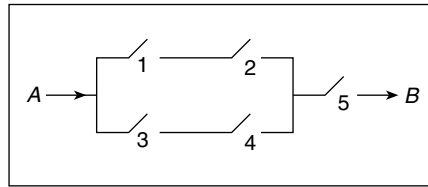
- (a) the probability that the husband earns less than \$50,000;
- (b) the conditional probability that the wife earns more than \$50,000 given that the husband earns more than this amount;

- (c) the conditional probability that the wife earns more than \$50,000 given that the husband earns less than this amount?
27. There are two local factories that produce microwaves. Each microwave produced at factory A is defective with probability .05, whereas each one produced at factory B is defective with probability .01. Suppose you purchase two microwaves that were produced at the same factory, which is equally likely to have been either factory A or factory B . If the first microwave that you check is defective, what is the conditional probability that the other one is also defective?
28. A red die, a blue die, and a yellow die (all six-sided) are rolled. We are interested in the probability that the number appearing on the blue die is less than that appearing on the yellow die which is less than that appearing on the red die. (That is, if B (R) [Y] is the number appearing on the blue (red) [yellow] die, then we are interested in $P(B < Y < R)$.)
- (a) What is the probability that no two of the dice land on the same number?
- (b) Given that no two of the dice land on the same number, what is the conditional probability that $B < Y < R$?
- (c) What is $P(B < Y < R)$?
- (d) If we regard the outcome of the experiment as the vector B, R, Y , how many outcomes are there in the sample space?
- (e) Without using the answer to (c), determine the number of outcomes that result in $B < Y < R$.
- (f) Use the results of parts (d) and (e) to verify your answer to part (c).
29. You ask your neighbor to water a sickly plant while you are on vacation. Without water it will die with probability .8; with water it will die with probability .15. You are 90 percent certain that your neighbor will remember to water the plant.
- (a) What is the probability that the plant will be alive when you return?
- (b) If it is dead, what is the probability your neighbor forgot to water it?
30. Two balls, each equally likely to be colored either red or blue, are put in an urn. At each stage one of the balls is randomly chosen, its color is noted, and it is then returned to the urn. If the first two balls chosen are colored red, what is the probability that
- (a) both balls in the urn are colored red;
- (b) the next ball chosen will be red?
31. A total of 600 of the 1,000 people in a retirement community classify themselves as Republicans, while the others classify themselves as Democrats. In a local election in which everyone voted, 60 Republicans voted for the Democratic candidate, and 50 Democrats voted for the Republican candidate. If a randomly chosen community member voted for the Republican, what is the probability that she or he is a Democrat?

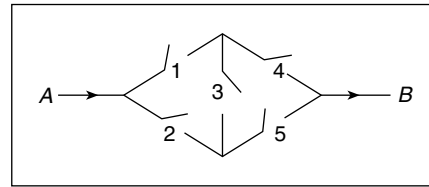
32. Each of 2 balls is painted black or gold and then placed in an urn. Suppose that each ball is colored black with probability $\frac{1}{2}$, and that these events are independent.
- (a) Suppose that you obtain information that the gold paint has been used (and thus at least one of the balls is painted gold). Compute the conditional probability that both balls are painted gold.
 - (b) Suppose, now, that the urn tips over and 1 ball falls out. It is painted gold. What is the probability that both balls are gold in this case? Explain.
33. Each of 2 cabinets identical in appearance has 2 drawers. Cabinet *A* contains a silver coin in each drawer, and cabinet *B* contains a silver coin in one of its drawers and a gold coin in the other. A cabinet is randomly selected, one of its drawers is opened, and a silver coin is found. What is the probability that there is a silver coin in the other drawer?
34. Prostate cancer is the most common type of cancer found in males. As an indicator of whether a male has prostate cancer, doctors often perform a test that measures the level of the PSA protein (prostate specific antigen) that is produced only by the prostate gland. Although higher PSA levels are indicative of cancer, the test is notoriously unreliable. Indeed, the probability that a noncancerous man will have an elevated PSA level is approximately .135, with this probability increasing to approximately .268 if the man does have cancer. If, based on other factors, a physician is 70 percent certain that a male has prostate cancer, what is the conditional probability that he has the cancer given that
- (a) the test indicates an elevated PSA level;
 - (b) the test does not indicate an elevated PSA level?
- Repeat the preceding, this time assuming that the physician initially believes there is a 30 percent chance the man has prostate cancer.
35. Suppose that an insurance company classifies people into one of three classes — good risks, average risks, and bad risks. Their records indicate that the probabilities that good, average, and bad risk persons will be involved in an accident over a 1-year span are, respectively, .05, .15, and .30. If 20 percent of the population are “good risks,” 50 percent are “average risks,” and 30 percent are “bad risks,” what proportion of people have accidents in a fixed year? If policy holder *A* had no accidents in 1987, what is the probability that he or she is a good (average) risk?
36. A pair of fair dice is rolled. Let *E* denote the event that the sum of the dice is equal to 7.
- (a) Show that *E* is independent of the event that the first die lands on 4.
 - (b) Show that *E* is independent of the event that the second die lands on 3.
37. The probability of the closing of the *i*th relay in the circuits shown is given by p_i , $i = 1, 2, 3, 4, 5$. If all relays function independently, what is the probability that a current flows between *A* and *B* for the respective circuits?



(a)



(b)



(c)

38. An engineering system consisting of n components is said to be a k -out-of- n system ($k \leq n$) if the system functions if and only if at least k of the n components function. Suppose that all components function independently of each other.
- If the i th component functions with probability P_i , $i = 1, 2, 3, 4$, compute the probability that a 2-out-of-4 system functions.
 - Repeat (a) for a 3-out-of-5 system.
39. Five independent flips of a fair coin are made. Find the probability that
- the first three flips are the same;
 - either the first three flips are the same, or the last three flips are the same;
 - there are at least two heads among the first three flips, and at least two tails among the last three flips.
40. Suppose that n independent trials, each of which results in any of the outcomes 0, 1, or 2, with respective probabilities .3, .5, and .2, are performed. Find the probability that both outcome 1 and outcome 2 occur at least once. (*Hint: Consider the complementary probability.*)
41. A parallel system functions whenever at least one of its components works. Consider a parallel system of n components, and suppose that each component independently works with probability $\frac{1}{2}$. Find the conditional probability that component 1 works, given that the system is functioning.
42. A certain organism possesses a pair of each of 5 different genes (which we will designate by the first 5 letters of the English alphabet). Each gene appears in 2 forms (which we designate by lowercase and capital letters). The capital letter will be assumed to be the dominant gene in the sense that if an organism possesses

the gene pair xX , then it will outwardly have the appearance of the X gene. For instance, if X stands for brown eyes and x for blue eyes, then an individual having either gene pair XX or xX will have brown eyes, whereas one having gene pair xx will be blue-eyed. The characteristic appearance of an organism is called its *phenotype*, whereas its genetic constitution is called its *genotype*. (Thus 2 organisms with respective genotypes aA, bB, cc, dD, ee and AA, BB, cc, DD, ee would have different genotypes but the same phenotype.) In a mating between 2 organisms each one contributes, at random, one of its gene pairs of each type. The 5 contributions of an organism (one of each of the 5 types) are assumed to be independent and are also independent of the contributions of its mate. In a mating between organisms having genotypes aA, bB, cC, dD, eE , and aa, bB, cc, Dd, ee , what is the probability that the progeny will (1) phenotypically, (2) genotypically resemble

- (a) the first parent;
 - (b) the second parent;
 - (c) either parent;
 - (d) neither parent?
43. 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 because he already knows that at least one of the two will go free. The jailer refuses to answer this question, pointing out that if A knew which of his fellow prisoners were to be set free, then his own probability of being executed would rise from $\frac{1}{3}$ to $\frac{1}{2}$ because he would then be one of two prisoners. What do you think of the jailer's reasoning?
44. Although both my parents have brown eyes, I have blue eyes. What is the probability that my sister has blue eyes? (As stated in Problem 42, an individual who receives a blue-eyed gene from each parent will have blue eyes, whereas one who receives one blue-eyed and one brown-eyed gene will have brown eyes.)
45. In a 7 game series played with two teams, the first team to win a total of 4 games is the winner. Suppose that each game played is independently won by team A with probability p .
- (a) Given that one team leads 3 to 0, what is the probability that it is team A that is leading?
 - (b) Given that one team leads 3 to 0, what is the probability that team wins the series?
 - (c) If $p = \frac{1}{2}$, what is the probability that the team that wins the first game wins the series?
46. Suppose that distinct integer values are written on each of 3 cards. Suppose you are to be offered these cards in a random order. When you are offered a card you

must immediately either accept it or reject it. If you accept a card, the process ends. If you reject a card, then the next card (if a card remains) is offered. If you reject the first two cards offered, then you must accept the final card.

- (a) If you plan to accept the first card offered, what is the probability that you will accept the highest valued card?
- (b) If you plan to reject the first card offered, and to then accept the second card if and only if its value is greater than the value of the first card, what is the probability that you will accept the highest valued card?

47. Let A , B , C be events such that $P(A) = .2$, $P(B) = .3$, $P(C) = .4$.

Find the probability that at least one of the events A and B occurs if

- (a) A and B are mutually exclusive;
- (b) A and B are independent.

Find the probability that all of the events A , B , C occur if

- (c) A , B , C are independent;
- (d) A , B , C are mutually exclusive.

48. Two percent of woman of age 45 who participate in routine screening have breast cancer. Ninety percent of those with breast cancer have positive mammographies. Ten percent of the women who do not have breast cancer will also have positive mammographies. Given a woman has a positive mammography, what is the probability she has breast cancer?

49. Twelve percent of all US households are in California. A total of 3.3 percent of all US households earn over 250,000 per year, while a total of 6.3 percent of California households earn over 250,000 per year. If a randomly chosen US household earns over 250,000 per year, what is the probability it is from California?

50. There is a 60 percent chance that the event A will occur. If A does not occur, there is a 10 percent chance that B will occur. What is the probability that at least one of the events A or B occur?

51. Suppose distinct values are written on each of three cards, which are then randomly given the designations A , B , and C . The values on cards A and B are then compared. What is the probability that the smaller of these values is also smaller than the value on card C ?