Chapter 2

Probability

Introduction

The development of the theory of probability was financed by seventeenth-century gamblers, who hired some of the leading mathematicians of the day to calculate the correct odds for certain games of chance. Later, people realized that scientific processes involve chance as well, and since then the methods of probability have been used to study the physical world.

Probability is now an extensive branch of mathematics. Many books are devoted to the subject, and many researchers have dedicated their professional careers to its further development. In this chapter we present an introduction to the ideas of probability that are most important to the study of statistics.

2.1 Basic Ideas

To make a systematic study of probability, we need some terminology. An **experiment** is a process that results in an outcome that cannot be predicted in advance with certainty. Tossing a coin, rolling a die, measuring the diameter of a bolt, weighing the contents of a box of cereal, and measuring the breaking strength of a length of fishing line are all examples of experiments. To discuss an experiment in probabilistic terms, we must specify its possible outcomes:

Definition

The set of all possible outcomes of an experiment is called the **sample space** for the experiment.

For tossing a coin, we can use the set $\{\text{Heads, Tails}\}$ as the sample space. For rolling a six-sided die, we can use the set $\{1, 2, 3, 4, 5, 6\}$. These sample spaces are finite. Some experiments have sample spaces with an infinite number of outcomes. For example, imagine that a punch with diameter 10 mm punches holes in sheet metal. Because of variations in the angle of the punch and slight movements in the sheet metal, the diameters of the holes vary between 10.0 and 10.2 mm. For the experiment of punching a hole, then, a reasonable sample space is the interval (10.0, 10.2), or in set notation, $\{x \mid 10.0 < x < 10.2\}$. This set obviously contains an infinite number of outcomes.

For many experiments, there are several sample spaces to choose from. For example, assume that a process manufactures steel pins whose lengths vary between 5.20 and 5.25 cm. An obvious choice for the sample space for the length of a pin is the set $\{x \mid 5.20 < x < 5.25\}$. However, if the object were simply to determine whether the pin was too short, too long, or within specification limits, a good choice for the sample space might be {too short, too long, within specifications}.

When discussing experiments, we are often interested in a particular subset of outcomes. For example, we might be interested in the probability that a die comes up an even number. The sample space for the experiment is $\{1, 2, 3, 4, 5, 6\}$, and coming up even corresponds to the subset $\{2, 4, 6\}$. In the hole punch example, we might be interested in the probability that a hole has a diameter less than 10.1 mm. This corresponds to the subset $\{x \mid 10.0 < x < 10.1\}$. There is a special name for a subset of a sample space:

Definition

A subset of a sample space is called an **event**.

Note that for any sample space, the empty set $\emptyset$ is an event, as is the entire sample space. A given event is said to have occurred if the outcome of the experiment is one of the outcomes in the event. For example, if a die comes up 2, the events $\{2, 4, 6\}$ and $\{1, 2, 3\}$ have both occurred, along with every other event that contains the outcome "2."

Example 2.1

An electrical engineer has on hand two boxes of resistors, with four resistors in each box. The resistors in the first box are labeled 10 Ω (ohms), but in fact their resistances are 9, 10, 11, and 12 Ω. The resistors in the second box are labeled 20 Ω, but in fact their resistances are 18, 19, 20, and 21 Ω. The engineer chooses one resistor from each box and determines the resistance of each.

Let A be the event that the first resistor has a resistance greater than 10, let B be the event that the second resistor has a resistance less than 19, and let C be the event that the sum of the resistances is equal to 28. Find a sample space for this experiment, and specify the subsets corresponding to the events A, B, and C.

Solution

A good sample space for this experiment is the set of ordered pairs in which the first component is the resistance of the first resistor and the second component is the resistance of the second resistor. We will denote this sample space by $\mathcal{S}$.

$$\mathcal{S} = \{(9, 18), (9, 19), (9, 20), (9, 21), (10, 18), (10, 19), (10, 20), (10, 21), (11, 18), (11, 19), (11, 20), (11, 21), (12, 18), (12, 19), (12, 20), (12, 21)\}$$

The events A, B, and C are given by

$A = \{(11, 18), (11, 19), (11, 20), (11, 21), (12, 18), (12, 19), (12, 20), (12, 21)\}$
$B = \{(9, 18), (10, 18), (11, 18), (12, 18)\}$
$C = \{(9, 19), (10, 18)\}$

Combining Events

We often construct events by combining simpler events. Because events are subsets of sample spaces, it is traditional to use the notation of sets to describe events constructed in this way. We review the necessary notation here.

- The **union** of two events A and B, denoted $A \cup B$, is the set of outcomes that belong either to A, to B, or to both. In words, $A \cup B$ means "A or B." Thus the event $A \cup B$ occurs whenever either A or B (or both) occurs.
- The **intersection** of two events A and B, denoted $A \cap B$, is the set of outcomes that belong both to A and to B. In words, $A \cap B$ means "A and B." Thus the event $A \cap B$ occurs whenever both A and B occur.
- The **complement** of an event A, denoted A^c, is the set of outcomes that do not belong to A. In words, A^c means "not A." Thus the event A^c occurs whenever A does *not* occur.

Events can be graphically illustrated with Venn diagrams. Figure 2.1 illustrates the events $A \cup B$, $A \cap B$, and $B \cap A^c$.

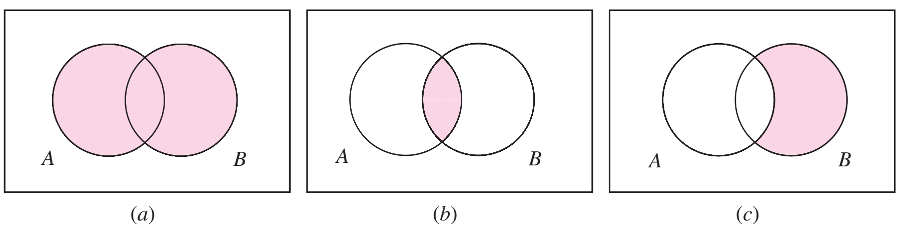


FIGURE 2.1 Venn diagrams illustrating various events: **(a)** $A \cup B$, **(b)** $A \cap B$, **(c)** $B \cap A^c$.

Example 2.2

Refer to Example 2.1. Find $B \cup C$ and $A \cap B^c$.

Solution

The event $B \cup C$ contains all the outcomes that belong either to B or to C, or to both. Therefore

$$B \cup C = \{(9, 18), (10, 18), (11, 18), (12, 18), (9, 19)\}$$

The event B^c contains those outcomes in the sample space that do not belong to B. It follows that the event $A \cap B^c$ contains the outcomes that belong to A and do not belong to B. Therefore

$$A \cap B^c = \{(11, 19), (11, 20), (11, 21), (12, 19), (12, 20), (12, 21)\}$$

Mutually Exclusive Events

There are some events that can never occur together. For example, it is impossible that a coin can come up both heads and tails, and it is impossible that a steel pin can be both too long and too short. Events like this are said to be **mutually exclusive**.

Definition

- The events A and B are said to be **mutually exclusive** if they have no outcomes in common.
- More generally, a collection of events $A_1, A_2, \ldots, A_n$ is said to be mutually exclusive if no two of them have any outcomes in common.

The Venn diagram in Figure 2.2 illustrates mutually exclusive events.

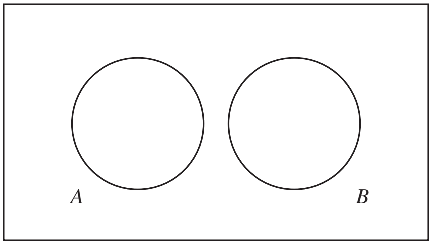


FIGURE 2.2 The events A and B are mutually exclusive.

Example 2.3

Refer to Example 2.1. If the experiment is performed, is it possible for events A and B both to occur? How about B and C? A and C? Which pair of events is mutually exclusive?

Solution

If the outcome is $(11, 18)$ or $(12, 18)$, then events A and B both occur. If the outcome is $(10, 18)$, then both B and C occur. It is impossible for A and C both to occur, because these events are mutually exclusive, having no outcomes in common.

Probabilities

Each event in a sample space has a **probability** of occurring. Intuitively, the probability is a quantitative measure of how likely the event is to occur. Formally speaking, there are several interpretations of probability; the one we shall adopt is that the probability of an event is the proportion of times the event would occur in the long run, if the experiment were to be repeated over and over again.

We often use the letter P to stand for probability. Thus when tossing a coin, the notation "$P(\text{heads}) = 1/2$" means that the probability that the coin lands heads is equal to $1/2$.

Summary

Given any experiment and any event A:

- The expression $P(A)$ denotes the probability that the event A occurs.
- $P(A)$ is the proportion of times that event A would occur in the long run, if the experiment were to be repeated over and over again.

In many situations, the only way to estimate the probability of an event is to repeat the experiment many times and determine the proportion of times that the event occurs. For example, if it is desired to estimate the probability that a printed circuit board manufactured by a certain process is defective, it is usually necessary to produce a number of boards and test them to determine the proportion that are defective. In some cases, probabilities can be determined through knowledge of the physical nature of an experiment. For example, if it is known that the shape of a die is nearly a perfect cube and that its mass is distributed nearly uniformly, it may be assumed that each of the six faces is equally likely to land upward when the die is rolled.

Once the probabilities of some events have been found through scientific knowledge or experience, the probabilities of other events can be computed mathematically. For example, if it has been estimated through experimentation that the probability that a printed circuit board is defective is 0.10, an estimate of the probability that a board is not defective can be calculated to be 0.90. As another example, assume that steel pins manufactured by a certain process can fail to meet a length specification either by being too short or too long. By measuring a large number of pins, it is estimated that the probability that a pin is too short is 0.02, and the probability that a pin is too long is 0.03. It can then be estimated that the probability that a pin fails to meet the specification is 0.05.

In practice, scientists and engineers estimate the probabilities of some events on the basis of scientific understanding and experience and then use mathematical rules to compute estimates of the probabilities of other events. In the rest of this section and in Section 2.2, we will explain some of these rules and show how to use them.

Axioms of Probability

The subject of probability is based on three commonsense rules, known as axioms. They are:

The Axioms of Probability

1. Let $\mathcal{S}$ be a sample space. Then $P(\mathcal{S}) = 1$.
2. For any event A, $0 \leq P(A) \leq 1$.
3. If A and B are mutually exclusive events, then $P(A \cup B) = P(A) + P(B)$. More generally, if $A_1, A_2, \ldots$ are mutually exclusive events, then $P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$.

With a little thought, it is easy to see that the three axioms do indeed agree with common sense. The first axiom says that the outcome of an experiment is always in the sample space. This is obvious, because by definition the sample space contains all the possible outcomes of the experiment. The second axiom says that the long-run frequency of any event is always between 0 and 100%. For an example illustrating the third axiom, we previously discussed a process that manufactures steel pins, in which the probability that a pin is too short is 0.02 and the probability that a pin is too long is 0.03. The third axiom says that the probability that the pin is either too short or too long is $0.02 + 0.03 = 0.05$.

We now present two simple rules that are helpful in computing probabilities. These rules are intuitively obvious, and they can also be proved from the axioms. Proofs are provided at the end of the section.

For any event A,

$$P(A^c) = 1 - P(A) \tag{2.1}$$

Let $\emptyset$ denote the empty set. Then

$$P(\emptyset) = 0 \tag{2.2}$$

Equation (2.1) says that the probability that an event does not occur is equal to 1 minus the probability that it does occur. For example, if there is a 40% chance of rain, there is a 60% chance that it does not rain. Equation (2.2) says that it is impossible for an experiment to have no outcome.

Example 2.4

A target on a test firing range consists of a bull's-eye with two concentric rings around it. A projectile is fired at the target. The probability that it hits the bull's-eye is 0.10, the probability that it hits the inner ring is 0.25, and the probability that it hits the outer ring is 0.45. What is the probability that the projectile hits the target? What is the probability that it misses the target?

Solution

Hitting the bull's-eye, hitting the inner ring, and hitting the outer ring are mutually exclusive events, since it is impossible for more than one of these events to occur. Therefore, using Axiom 3,

$$
\begin{aligned}
P(\text{hits target}) &= P(\text{bull's-eye}) + P(\text{inner ring}) + P(\text{outer ring}) \\
&= 0.10 + 0.25 + 0.45 \\
&= 0.80
\end{aligned}
$$

We can now compute the probability that the projectile misses the target by using Equation (2.1):

$$
\begin{aligned}
P(\text{misses target}) &= 1 - P(\text{hits target}) \\
&= 1 - 0.80 \\
&= 0.20
\end{aligned}
$$

Example 2.5

The following table presents probabilities for the number of times that a certain computer system will crash in the course of a week. Let A be the event that there are more than two crashes during the week, and let B be the event that the system crashes at least once. Find a sample space. Then find the subsets of the sample space that correspond to the events A and B. Then find $P(A)$ and $P(B)$.

Number of Crashes	Probability
0	0.60
1	0.30
2	0.05
3	0.04
4	0.01

Solution

A sample space for the experiment is the set $\{0, 1, 2, 3, 4\}$. The events are $A = \{3, 4\}$ and $B = \{1, 2, 3, 4\}$. To find $P(A)$, notice that A is the event that either 3 crashes happen or 4 crashes happen. The events "3 crashes happen" and "4 crashes happen" are mutually exclusive. Therefore, using Axiom 3, we conclude that

$$
\begin{aligned}
P(A) &= P(\text{3 crashes happen or 4 crashes happen}) \\
&= P(\text{3 crashes happen}) + P(\text{4 crashes happen}) \\
&= 0.04 + 0.01 \\
&= 0.05
\end{aligned}
$$

We will compute $P(B)$ in two ways. First, note that B^c is the event that no crashes happen. Therefore, using Equation (2.1),

$$
\begin{aligned}
P(B) &= 1 - P(B^c) \\
&= 1 - P(\text{0 crashes happen}) \\
&= 1 - 0.60 \\
&= 0.40
\end{aligned}
$$

For a second way to compute $P(B)$, note that B is the event that 1 crash happens or 2 crashes happen or 3 crashes happen or 4 crashes happen. These events are mutually exclusive. Therefore, using Axiom 3, we conclude that

$$\begin{aligned} P(B) &= P(1 \text{ crash}) + P(2 \text{ crashes}) + P(3 \text{ crashes}) + P(4 \text{ crashes}) \\ &= 0.30 + 0.05 + 0.04 + 0.01 \\ &= 0.40 \end{aligned}$$

In Example 2.5, we computed the probabilities of the events $A = \{3, 4\}$ and $B = \{1, 2, 3, 4\}$ by summing the probabilities of the outcomes in each of the events: $P(A) = P(3) + P(4)$ and $P(B) = P(1) + P(2) + P(3) + P(4)$. This method works in general. Since any two outcomes in a sample space are mutually exclusive, the probability of any event that contains a finite number of outcomes can be found by summing the probabilities of the outcomes that comprise the event.

> If A is an event containing outcomes $O_1, \ldots, O_n$, that is, if $A = \{O_1, \ldots, O_n\}$, then
>
> $$P(A) = P(O_1) + P(O_2) + \cdots + P(O_n) \qquad (2.3)$$

Sample Spaces with Equally Likely Outcomes

For some experiments, a sample space can be constructed in which all the outcomes are equally likely. A simple example is the roll of a fair die, in which the sample space is $\{1, 2, 3, 4, 5, 6\}$ and each of these outcomes has probability $1/6$. Another type of experiment that results in equally likely outcomes is the random selection of an item from a population of items. The items in the population can be thought of as the outcomes in a sample space, and each item is equally likely to be selected.

> A population from which an item is sampled at random can be thought of as a sample space with equally likely outcomes.

If a sample space contains N equally likely outcomes, the probability of each outcome is $1/N$. This is so, because the probability of the whole sample space must be 1, and this probability is equally divided among the N outcomes. If A is an event that contains k outcomes, then $P(A)$ can be found by summing the probabilities of the k outcomes, so $P(A) = k/N$.

> If $\mathcal{S}$ is a sample space containing N equally likely outcomes, and if A is an event containing k outcomes, then
>
> $$P(A) = \frac{k}{N} \qquad (2.4)$$

Example 2.6

An extrusion die is used to produce aluminum rods. Specifications are given for the length and the diameter of the rods. For each rod, the length is classified as too short, too long, or OK, and the diameter is classified as too thin, too thick, or OK. In a population of 1000 rods, the number of rods in each class is as follows:

	Diameter		
Length	**Too Thin**	**OK**	**Too Thick**
Too Short	10	3	5
OK	38	900	4
Too Long	2	25	13

A rod is sampled at random from this population. What is the probability that it is too short?

Solution

We can think of each of the 1000 rods as an outcome in a sample space. Each of the 1000 outcomes is equally likely. We'll solve the problem by counting the number of outcomes that correspond to the event. The number of rods that are too short is $10 + 3 + 5 = 18$. Since the total number of rods is 1000,

$$P(\text{too short}) = \frac{18}{1000}$$

The Addition Rule

If A and B are mutually exclusive events, then $P(A \cup B) = P(A) + P(B)$. This rule can be generalized to cover the case where A and B are not mutually exclusive. Example 2.7 illustrates the reasoning.

Example 2.7

Refer to Example 2.6. If a rod is sampled at random, what is the probability that it is either too short or too thick?

Solution

First we'll solve this problem by counting the number of outcomes that correspond to the event. In the following table the numbers of rods that are too thick are circled, and the numbers of rods that are too short have rectangles around them. Note that there are 5 rods that are both too short and too thick.

	Diameter		
Length	**Too Thin**	**OK**	**Too Thick**
Too Short	$\boxed{10}$	$\boxed{3}$	$\boxed{5}$
OK	38	900	4
Too Long	2	25	13

Of the 1000 outcomes, the number that are either too short or too thick is $10 + 3 + 5 + 4 + 13 = 35$. Therefore

$$P(\text{too short or too thick}) = \frac{35}{1000}$$

Now we will solve the problem in a way that leads to a more general method. In the sample space, there are $10 + 3 + 5 = 18$ rods that are too short and $5 + 4 + 13 = 22$ rods that are too thick. But if we try to find the number of rods that are either too short or too thick by adding $18 + 22$, we get too large a number (40 instead of 35). The reason is that there are five rods that are both too short and too thick, and these are counted twice. We can still solve the problem by adding 18 and 22, but we must then subtract 5 to correct for the double counting.

We restate this reasoning, using probabilities:

$$P(\text{too short}) = \frac{18}{1000}, \quad P(\text{too thick}) = \frac{22}{1000}, \quad P(\text{too short and too thick}) = \frac{5}{1000}$$

$$\begin{aligned} P(\text{too short or too thick}) &= P(\text{too short}) + P(\text{too thick}) - P(\text{too short and too thick}) \\ &= \frac{18}{1000} + \frac{22}{1000} - \frac{5}{1000} \\ &= \frac{35}{1000} \end{aligned}$$

The method of Example 2.7 holds for any two events in any sample space. In general, to find the probability that either of two events occurs, add the probabilities of the events and then subtract the probability that they both occur.

Summary

Let A and B be any events. Then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{2.5}$$

A proof of this result, based on the axioms, is provided at the end of this section. Note that if A and B are mutually exclusive, then $P(A \cap B) = 0$, so Equation (2.5) reduces to Axiom 3 in this case.

Example 2.8

In a process that manufactures aluminum cans, the probability that a can has a flaw on its side is 0.02, the probability that a can has a flaw on the top is 0.03, and the probability that a can has a flaw on both the side and the top is 0.01. What is the probability that a randomly chosen can has a flaw? What is the probability that it has no flaw?

Solution

We are given that $P(\text{flaw on side}) = 0.02$, $P(\text{flaw on top}) = 0.03$, and $P(\text{flaw on side and flaw on top}) = 0.01$. Now $P(\text{flaw}) = P(\text{flaw on side or flaw on top})$. Using Equation (2.5),

$$\begin{aligned} P(\text{flaw on side or flaw on top}) &= P(\text{flaw on side}) + P(\text{flaw on top}) \\ &\quad - P(\text{flaw on side and flaw on top}) \\ &= 0.02 + 0.03 - 0.01 \\ &= 0.04 \end{aligned}$$

To find the probability that a can has no flaw, we compute

$$P(\text{no flaw}) = 1 - P(\text{flaw}) = 1 - 0.04 = 0.96$$

Venn diagrams can sometimes be useful in computing probabilities by showing how to express an event as the union of disjoint events. Example 2.9 illustrates the method.

Example 2.9

Refer to Example 2.8. What is the probability that a can has a flaw on the top but not on the side?

Solution

Let A be the event that a can has a flaw on the top and let B be the event that a can has a flaw on the side. We need to find $P(A \cap B^c)$. The following Venn diagram (Figure 2.3) shows that $A \cap B$ and $A \cap B^c$ are mutually exclusive, so that

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

We know that $P(A) = 0.03$ and $P(A \cap B) = 0.01$. Therefore $0.03 = 0.01 + P(A \cap B^c)$, so $P(A \cap B^c) = 0.02$.

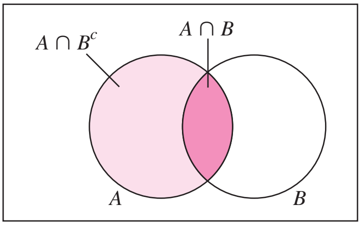


FIGURE 2.3 The events $A \cap B$ and $A \cap B^c$ are mutually exclusive, and their union is the event A.

Proof that $P(A^c) = 1 - P(A)$

Let $\mathcal{S}$ be a sample space and let A be an event. Then A and A^c are mutually exclusive, so by Axiom 3,

$$P(A \cup A^c) = P(A) + P(A^c)$$

But $A \cup A^c = \mathcal{S}$, and by Axiom 1, $P(\mathcal{S}) = 1$. Therefore

$$P(A \cup A^c) = P(\mathcal{S}) = 1$$

It follows that $P(A) + P(A^c) = 1$, so $P(A^c) = 1 - P(A)$.

Proof that $P(\emptyset) = 0$

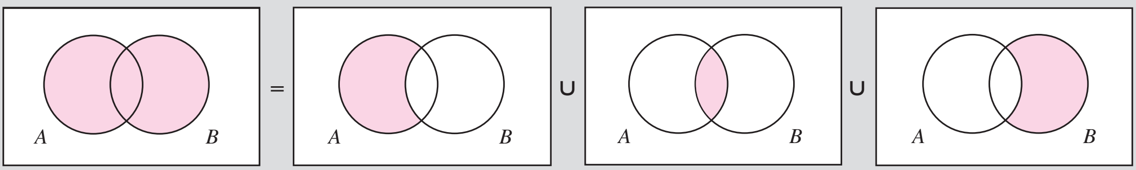
Let $\mathcal{S}$ be a sample space. Then $\emptyset = \mathcal{S}^c$. Therefore $P(\emptyset) = 1 - P(\mathcal{S}) = 1 - 1 = 0$.

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

Let A and B be any two events. The key to the proof is to write $A \cup B$ as the union of three mutually exclusive events: $A \cap B^c$, $A \cap B$, and $A^c \cap B$.

$$A \cup B = (A \cap B^c) \cup (A \cap B) \cup (A^c \cap B) \tag{2.6}$$

The following figure illustrates Equation (2.6).



By Axiom 3,

$$P(A \cup B) = P(A \cap B^c) + P(A \cap B) + P(A^c \cap B) \tag{2.7}$$

Now $A = (A \cap B^c) \cup (A \cap B)$, and $B = (A^c \cap B) \cup (A \cap B)$. Therefore

$$P(A) = P(A \cap B^c) + P(A \cap B) \tag{2.8}$$

and

$$P(B) = P(A^c \cap B) + P(A \cap B) \tag{2.9}$$

Summing Equations (2.8) and (2.9) yields

$$P(A) + P(B) = P(A \cap B^c) + P(A^c \cap B) + 2P(A \cap B) \tag{2.10}$$

Comparing Equations (2.10) and (2.7) shows that

$$P(A) + P(B) = P(A \cup B) + P(A \cap B) \tag{2.11}$$

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

Exercises for Section 2.1

1. The probability that a bearing fails during the first month of use is 0.12. What is the probability that it does not fail during the first month?

2. A die (six faces) has the number 1 painted on three of its faces, the number 2 painted on two of its faces, and the number 3 painted on one face. Assume that each face is equally likely to come up.
   a. Find a sample space for this experiment.
   b. Find P(odd number).
   c. If the die were loaded so that the face with the 3 on it were twice as likely to come up as each of the other five faces, would this change the sample space? Explain.
   d. If the die were loaded so that the face with the 3 on it were twice as likely to come up as each of the other five faces, would this change the value of P(odd number)? Explain.

3. A section of an exam contains four True-False questions. A completed exam paper is selected at random, and the four answers are recorded.
   a. List all 16 outcomes in the sample space.
   b. Assuming the outcomes to be equally likely, find the probability that all the answers are the same.
   c. Assuming the outcomes to be equally likely, find the probability that exactly one of the four answers is "True."
   d. Assuming the outcomes to be equally likely, find the probability that at most one of the four answers is "True."

4. A commuter passes through three traffic lights on the way to work. Each light is either red, yellow, or green. An experiment consists of observing the colors of the three lights.
   a. List the 27 outcomes in the sample space.
   b. Let A be the event that all the colors are the same. List the outcomes in A.
   c. Let B be the event that all the colors are different. List the outcomes in B.
   d. Let C be the event that at least two lights are green. List the outcomes in C.
   e. List the outcomes in $A \cap C$.
   f. List the outcomes in $A \cup B$.
   g. List the outcomes in $A \cap C^c$.
   h. List the outcomes in $A^c \cap C$.
   i. Are events A and C mutually exclusive? Explain.
   j. Are events B and C mutually exclusive? Explain.

5. A box contains four bolts. Two of them, labeled #1 and #2, are 5 mm in diameter, and two of them, labeled #3 and #4, are 7 mm in diameter. Bolts are randomly selected until a 5 mm bolt is obtained. The outcomes are the sequences of bolts that can be selected. So one outcome is 1, and another is 342.
   a. List all the possible outcomes.
   b. Let A be the event that only one bolt is selected. List the outcomes in A.
   c. Let B be the event that three bolts are selected. List the outcomes in B.
   d. Let C be the event that bolt #2 is selected. List the outcomes in C.
   e. Let D be the event that bolt #4 is not selected. List the outcomes in D.
   f. Let E be the event that bolt #1 is selected. Are A and E mutually exclusive? How about B and E, C and E, D and E?

6. Two bolts are randomly selected from the box described in Exercise 5.
   a. List the equally likely outcomes.
   b. What is the probability that both bolts are 7 mm?
   c. What is the probability that one bolt is 5 mm and the other is 7 mm?

7. In a survey of households with television sets, the proportion of television sets in various types of rooms was

Room	Proportion of TV Sets
Bedroom	0.37
Living Room	0.26
Den	0.22
Basement	0.12
Kitchen	0.02
Bathroom	0.01

   a. What is the probability that a TV set is located in a living room or den?

b. What is the probability that a TV set is not located in a bedroom?

8. An automobile insurance company divides customers into three categories, good risks, medium risks, and poor risks. Assume that 70% of the customers are good risks, 20% are medium risks, and 10% are poor risks. As part of an audit, one customer is chosen at random.

a. What is the probability that the customer is a good risk?

b. What is the probability that the customer is not a poor risk?

9. Among the cast aluminum parts manufactured on a certain day, 80% were flawless, 15% had only minor flaws, and 5% had major flaws. Find the probability that a randomly chosen part

a. has a flaw (major or minor).

b. has no major flaw.

10. An item manufactured by a certain process has probability 0.10 of being defective. True or false:

a. If a sample of 100 items is drawn, exactly 10 of them will be defective.

b. If a sample of 100 items is drawn, the number of defectives is likely to be close to 10, but not exactly equal to 10.

c. As more and more items are sampled, the proportion of defective items will approach 10%.

11. A quality-control engineer samples 100 items manufactured by a certain process, and finds that 15 of them are defective. True or false:

a. The probability that an item produced by this process is defective is 0.15.

b. The probability that an item produced by this process is defective is likely to be close to 0.15, but not exactly equal to 0.15.

12. Let V be the event that a computer contains a virus, and let W be the event that a computer contains a worm. Suppose $P(V) = 0.15$, $P(W) = 0.05$, and $P(V \cup W) = 0.17$.

a. Find the probability that the computer contains both a virus and a worm.

b. Find the probability that the computer contains neither a virus nor a worm.

c. Find the probability that the computer contains a virus but not a worm.

13. Let S be the event that a randomly selected college student has taken a statistics course, and let C be the event that the same student has taken a chemistry course. Suppose $P(S) = 0.4$, $P(C) = 0.3$, and $P(S \cap C) = 0.2$.

a. Find the probability that a student has taken statistics, chemistry, or both.

b. Find the probability that a student has taken neither statistics nor chemistry.

c. Find the probability that a student has taken statistics but not chemistry.

14. Inspector A visually inspected 1000 ceramic bowls for surface flaws and found flaws in 37 of them. Inspector B inspected the same bowls and found flaws in 43 of them. A total of 948 bowls were found to be flawless by both inspectors. One of the 1000 bowls is selected at random.

a. Find the probability that a flaw was found in this bowl by at least one of the two inspectors.

b. Find the probability that flaws were found in this bowl by both inspectors.

c. Find the probability that a flaw was found by inspector A but not by inspector B.

15. All the fourth-graders in a certain elementary school took a standardized test. A total of 85% of the students were found to be proficient in reading, 78% were found to be proficient in mathematics, and 65% were found to be proficient in both reading and mathematics. A student is chosen at random.

a. What is the probability that the student is proficient in mathematics but not in reading?

b. What is the probability that the student is proficient in reading but not in mathematics?

c. What is the probability that the student is proficient in neither reading nor mathematics?

16. A system contains two components, A and B. The system will function so long as either A or B functions. The probability that A functions is 0.95, the probability that B functions is 0.90, and the probability that both function is 0.88. What is the probability that the system functions?

17. A system contains two components, A and B. The system will function only if both components function. The probability that A functions is 0.98, the probability that B functions is 0.95, and the probability that either A or B functions is 0.99. What is the probability that the system functions?

18. Human blood may contain either or both of two antigens, A and B. Blood that contains only the A antigen is called type A, blood that contains only the B antigen is called type B, blood that contains both antigens is called type AB, and blood that contains neither antigen is called type O. At a certain blood bank, 35% of the blood donors have type A blood, 10% have type B, and 5% have type AB.

a. What is the probability that a randomly chosen blood donor is type O?

b. A recipient with type A blood may safely receive blood from a donor whose blood does not contain the B antigen. What is the probability that a randomly chosen blood donor may donate to a recipient with type A blood?

19. True or false: If A and B are mutually exclusive,

a. $P(A \cup B) = 0$

b. $P(A \cap B) = 0$

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

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

20. A flywheel is attached to a crankshaft by 12 bolts, numbered 1 through 12. Each bolt is checked to determine whether it is torqued correctly. Let A be the event that all the bolts are torqued correctly, let B be the event that the #3 bolt is not torqued correctly, let C be the event that exactly one bolt is not torqued correctly, and let D be the event that bolts #5 and #8 are torqued correctly. State whether each of the following pairs of events is mutually exclusive.

a. A and B

b. B and D

c. C and D

d. B and C

2.2 Counting Methods

When computing probabilities, it is sometimes necessary to determine the number of outcomes in a sample space. In this section we will describe several methods for doing this. The basic rule, which we will call **the fundamental principle of counting**, is presented by means of Example 2.10.

Example 2.10

A certain make of automobile is available in any of three colors: red, blue, or green, and comes with either a large or small engine. In how many ways can a buyer choose a car?

Solution

There are three choices of color and two choices of engine. A complete list of choices is written in the following 3×2 table. The total number of choices is $(3)(2) = 6$.

	Red	Blue	Green
Large	Red, Large	Blue, Large	Green, Large
Small	Red, Small	Blue, Small	Green, Small

To generalize Example 2.10, if there are n_1 choices of color and n_2 choices of engine, a complete list of choices can be written in an $n_1 \times n_2$ table, so the total number of choices is $n_1 n_2$.

> If an operation can be performed in n_1 ways, and if for each of these ways a second operation can be performed in n_2 ways, then the total number of ways to perform the two operations is $n_1 n_2$.

The fundamental principle of counting states that this reasoning can be extended to any number of operations.

> **The Fundamental Principle of Counting**
> Assume that k operations are to be performed. If there are n_1 ways to perform the first operation, and if for each of these ways there are n_2 ways to perform the second operation, and if for each choice of ways to perform the first two operations there are n_3 ways to perform the third operation, and so on, then the total number of ways to perform the sequence of k operations is $n_1 n_2 \cdots n_k$.

Example 2.11

When ordering a certain type of computer, there are 3 choices of hard drive, 4 choices for the amount of memory, 2 choices of video card, and 3 choices of monitor. In how many ways can a computer be ordered?

Solution

The total number of ways to order a computer is $(3)(4)(2)(3) = 72$.

Permutations

A **permutation** is an ordering of a collection of objects. For example, there are six permutations of the letters A, B, C: ABC, ACB, BAC, BCA, CAB, and CBA. With only three objects, it is easy to determine the number of permutations just by listing them all. But with a large number of objects this would not be feasible. The fundamental principle of counting can be used to determine the number of permutations of any set of objects. For example, we can determine the number of permutations of a set of three objects as follows. There are 3 choices for the object to place first. After that choice is made, there are 2 choices remaining for the object to place second. Then there is 1 choice left for the object to place last. Therefore, the total number of ways to order three objects is $(3)(2)(1) = 6$. This reasoning can be generalized. The number of permutations of a collection of n objects is

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

This is the product of the integers from 1 to n. This product can be written with the symbol $n!$, read "n factorial."

Definition

For any positive integer n, $n! = n(n-1)(n-2)\cdots(3)(2)(1)$.

Also, we define $0! = 1$.

The number of permutations of n objects is $n!$.

Example 2.12

Five people stand in line at a movie theater. Into how many different orders can they be arranged?

Solution

The number of permutations of a collection of five people is $5! = (5)(4)(3)(2)(1) = 120$.

Sometimes we are interested in counting the number of permutations of subsets of a certain size chosen from a larger set. This is illustrated in Example 2.13.

Example 2.13

Five lifeguards are available for duty one Saturday afternoon. There are three lifeguard stations. In how many ways can three lifeguards be chosen and ordered among the stations?

Solution

We use the fundamental principle of counting. There are 5 ways to choose a lifeguard to occupy the first station, then 4 ways to choose a lifeguard to occupy the second station, and finally 3 ways to choose a lifeguard to occupy the third station. The total number of permutations of three lifeguards chosen from 5 is therefore $(5)(4)(3) = 60$.

The reasoning used to solve Example 2.13 can be generalized. The number of permutations of k objects chosen from a group of n objects is

$$(n)(n-1)\cdots(n-k+1)$$

This expression can be simplified by using factorial notation:

$$\begin{aligned}(n)(n-1)\cdots(n-k+1) &= \frac{n(n-1)\cdots(n-k+1)(n-k)(n-k-1)\cdots(3)(2)(1)}{(n-k)(n-k-1)\cdots(3)(2)(1)} \\ &= \frac{n!}{(n-k)!}\end{aligned}$$

Summary

The number of permutations of k objects chosen from a group of n objects is

$$\frac{n!}{(n-k)!}$$

Combinations

In some cases, when choosing a set of objects from a larger set, we don't care about the ordering of the chosen objects; we care only which objects are chosen. For example, we may not care which lifeguard occupies which station; we might care only which three lifeguards are chosen. Each distinct group of objects that can be selected, without regard to order, is called a **combination**. We will now show how to determine the number of combinations of k objects chosen from a set of n objects. We will illustrate the reasoning with the result of Example 2.13. In that example, we showed that there are 60 permutations of three objects chosen from five. Denoting the objects A, B, C, D, E, Figure 2.4 presents a list of all 60 permutations.

ABC	ABD	ABE	ACD	ACE	ADE	BCD	BCE	BDE	CDE
ACB	ADB	AEB	ADC	AEC	AED	BDC	BEC	BED	CED
BAC	BAD	BAE	CAD	CAE	DAE	CBD	CBE	DBE	DCE
BCA	BDA	BEA	CDA	CEA	DEA	CDB	CEB	DEB	DEC
CAB	DAB	EAB	DAC	EAC	EAD	DBC	EBC	EBD	ECD
CBA	DBA	EBA	DCA	ECA	EDA	DCB	ECB	EDB	EDC

FIGURE 2.4 The 60 permutations of three objects chosen from five.

The 60 permutations in Figure 2.4 are arranged in 10 columns of 6 permutations each. Within each column, the three objects are the same, and the column contains the six different permutations of those three objects. Therefore, each column represents a distinct combination of three objects chosen from five, and there are 10 such combinations. Figure 2.4 thus shows that the number of combinations of three objects chosen from five can be found by dividing the number of permutations of three objects chosen from five, which is $5!/(5-3)!$, by the number of permutations of three objects, which is 3! In summary, the number of combinations of three objects chosen from five is $\frac{5!}{3!(5-3)!}$.

The number of combinations of k objects chosen from n is often denoted by the symbol $\binom{n}{k}$. The reasoning used to derive the number of combinations of three objects chosen from five can be generalized to derive an expression for $\binom{n}{k}$.

Summary

The number of combinations of k objects chosen from a group of n objects is

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{2.12}$$

Example 2.14

At a certain event, 30 people attend, and 5 will be chosen at random to receive door prizes. The prizes are all the same, so the order in which the people are chosen does not matter. How many different groups of five people can be chosen?

Solution

Since the order of the five chosen people does not matter, we need to compute the number of combinations of 5 chosen from 30. This is

$$\begin{aligned}\binom{30}{5} &= \frac{30!}{5!25!} \\ &= \frac{(30)(29)(28)(27)(26)}{(5)(4)(3)(2)(1)} \\ &= 142{,}506\end{aligned}$$

Choosing a combination of k objects from a set of n divides the n objects into two subsets: the k that were chosen and the $n - k$ that were not chosen. Sometimes a set is to be divided up into more than two subsets. For example, assume that in a class of 12 students, a project is assigned in which the students will work in groups. Three groups are to be formed, consisting of five, four, and three students. We can calculate the number of ways in which the groups can be formed as follows. We consider the process of dividing the class into three groups as a sequence of two operations. The first operation is to select a combination of 5 students to comprise the group of 5. The second operation is to select a combination of 4 students from the remaining 7, to comprise the group of 4. The group of 3 will then automatically consist of the 3 students who are left.

The number of ways to perform the first operation is

$$\binom{12}{5} = \frac{12!}{5!7!}$$

After the first operation has been performed, the number of ways to perform the second operation is

$$\binom{7}{4} = \frac{7!}{4!3!}$$

The total number of ways to perform the sequence of two operations is therefore

$$\frac{12!}{5!7!}\,\frac{7!}{4!3!} = \frac{12!}{5!4!3!} = 27{,}720$$

Notice that the numerator in the final answer is the factorial of the total group size, while the denominator is the product of the factorials of the sizes of the groups chosen from it. This holds in general.

Summary

The number of ways of dividing a group of n objects into groups of $k_1, \ldots, k_r$ objects, where $k_1 + \cdots + k_r = n$, is

$$\frac{n!}{k_1! \cdots k_r!} \tag{2.13}$$

Example 2.15

A die is rolled 20 times. Given that three of the rolls came up 1, five came up 2, four came up 3, two came up 4, three came up 5, and three came up 6, how many different arrangements of the outcomes are there?

Solution

There are 20 outcomes. They are divided into six groups, namely, the group of three outcomes that came up 1, the group of five outcomes that came up 2, and so on. The number of ways to divide the 20 outcomes into six groups of the specified sizes is

$$\frac{20!}{3!5!4!2!3!3!} = 1.955 \times 10^{12}$$

When a sample space consists of equally likely outcomes, the probability of an event can be found by dividing the number of outcomes in the event by the total number of outcomes in the sample space. Counting rules can sometimes be used to compute these numbers, as the following example illustrates:

Example 2.16

A box of bolts contains 8 thick bolts, 5 medium bolts, and 3 thin bolts. A box of nuts contains 6 that fit the thick bolts, 4 that fit the medium bolts, and 2 that fit the thin bolts. One bolt and one nut are chosen at random. What is the probability that the nut fits the bolt?

Solution

The sample space consists of all pairs of nuts and bolts, and each pair is equally likely to be chosen. The event that the nut fits the bolt corresponds to the set of all matching pairs of nuts and bolts. Therefore

$$P(\text{nut fits bolt}) = \frac{\text{number of matching pairs of nuts and bolts}}{\text{number of pairs of nuts and bolts}}$$

There are $6 + 4 + 2 = 12$ nuts, and $8 + 5 + 3 = 16$ bolts. Therefore

$$\text{Number of pairs of nuts and bolts} = (12)(16) = 192$$

The number of matching pairs is found by summing the number of pairs of thick nuts and bolts, the number of pairs of medium nuts and bolts, and the number of pairs of thin nuts and bolts. These numbers are

$$\text{Number of pairs of thick nuts and bolts} = (6)(8) = 48$$
$$\text{Number of pairs of medium nuts and bolts} = (4)(5) = 20$$
$$\text{Number of pairs of thin nuts and bolts} = (2)(3) = 6$$

Therefore

$$P(\text{nut fits bolt}) = \frac{48 + 20 + 6}{192} = 0.3854$$

Exercises for Section 2.2

1. DNA molecules consist of chemically linked sequences of the bases adenine, guanine, cytosine, and thymine, denoted A, G, C, and T. A sequence of three bases is called a *codon*. A base may appear more than once in a codon.

a. How many different codons are there?

b. The bases A and G are *purines*, while C and T are *pyrimidines*. How many codons are there whose first and third bases are purines and whose second base is a pyrimidine?

c. How many codons consist of three different bases?

2. A chemical engineer is designing an experiment to determine the effect of temperature, stirring rate, and type of catalyst on the yield of a certain reaction. She wants to study five different reaction temperatures, two different stirring rates, and four different catalysts. If each run of the experiment involves a choice of one temperature, one stirring rate, and one catalyst, how many different runs are possible?

3. The article "Improved Bioequivalence Assessment of Topical Dermatological Drug Products Using Dermatopharmacokinetics" (B. N'Dri-Stempfer, W. Navidi, et al., *Pharmaceutical Research*, 2009:316–328) describes a study in which a new type of ointment was applied to forearms of volunteers to study the rates of absorption into the skin. Eight locations on the forearm were designated for ointment application. The new ointment was applied to four locations, and a control was applied to the other four. How many different choices were there for the four locations to apply the new ointment?

4. A group of 10 people have gotten together to play basketball. They will begin by dividing themselves into two teams of 5 players each. One team will wear red uniforms and the other will wear blue uniforms. In how many ways can this be done?

5. In horse racing, one can make a trifecta bet by specifying which horse will come in first, which will come in second, and which will come in third, in the correct order. One can make a box trifecta bet by specifying which three horses will come in first, second, and third, without specifying the order.

a. In an eight-horse field, how many different ways can one make a trifecta bet?

b. In an eight-horse field, how many different ways can one make a box trifecta bet?

6. A committee of eight people must choose a president, a vice-president, and a secretary. In how many ways can this be done?

7. A test consists of 15 questions. Ten are true-false questions, and five are multiple-choice questions that have four choices each. A student must select an answer for each question. In how many ways can this be done?

8. In a certain state, license plates consist of three letters followed by three numbers.

a. How many different license plates can be made?

b. How many different license plates can be made in which no letter or number appears more than once?

c. A license plate is chosen at random. What is the probability that no letter or number appears more than once?

9. A computer password consists of eight characters.

a. How many different passwords are possible if each character may be any lowercase letter or digit?

b. How many different passwords are possible if each character may be any lowercase letter or digit, and at least one character must be a digit?

c. A computer system requires that passwords contain at least one digit. If eight characters are generated at random, and each is equally likely to be any of the 26 letters or 10 digits, what is the probability that a valid password will be generated?

10. A company has hired 15 new employees, and must assign 6 to the day shift, 5 to the graveyard shift, and 4 to the night shift. In how many ways can the assignment be made?

11. One drawer in a dresser contains 8 blue socks and 6 white socks. A second drawer contains 4 blue socks and 2 white socks. One sock is chosen from each drawer. What is the probability that they match?

12. A drawer contains 6 red socks, 4 green socks, and 2 black socks. Two socks are chosen at random. What is the probability that they match?

2.3 Conditional Probability and Independence

A sample space contains all the possible outcomes of an experiment. Sometimes we obtain some additional information about an experiment that tells us that the outcome comes from a certain part of the sample space. In this case, the probability of an event is based on the outcomes in that part of the sample space. A probability that is based on a part of a sample space is called a **conditional probability**. We explore this idea through some examples.

In Example 2.6 (in Section 2.1) we discussed a population of 1000 aluminum rods. For each rod, the length is classified as too short, too long, or OK, and the diameter is classified as too thin, too thick, or OK. These 1000 rods form a sample space in which each rod is equally likely to be sampled. The number of rods in each category is presented in Table 2.1. Of the 1000 rods, 928 meet the diameter specification. Therefore, if a rod is sampled, $P(\text{diameter OK}) = 928/1000 = 0.928$. This probability is called the **unconditional probability**, since it is based on the entire sample space. Now assume that a rod is sampled, and its length is measured and found to meet the specification. What is the probability that the diameter also meets the specification? The key to computing this probability is to realize that knowledge that the length meets the specification reduces the sample space from which the rod is drawn. Table 2.2 (page 70) presents this idea. Once we know that the length specification is met, we know that the rod will be one of the 942 rods in the sample space presented in Table 2.2.

TABLE 2.1 Sample space containing 1000 aluminum rods

	Diameter		
Length	**Too Thin**	**OK**	**Too Thick**
Too Short	10	3	5
OK	38	900	4
Too Long	2	25	13

TABLE 2.2 Reduced sample space containing 942 aluminum rods that meet the length specification

	Diameter		
Length	**Too Thin**	**OK**	**Too Thick**
Too Short	—	—	—
OK	38	900	4
Too Long	—	—	—

Of the 942 rods in this sample space, 900 of them meet the diameter specification. Therefore, if we know that the rod meets the length specification, the probability that the rod meets the diameter specification is 900/942. We say that the **conditional probability** that the rod meets the diameter specification **given** that it meets the length specification is equal to 900/942, and we write $P(\text{diameter OK} \mid \text{length OK}) = 900/942 = 0.955$. Note that the conditional probability $P(\text{diameter OK} \mid \text{length OK})$ differs from the unconditional probability $P(\text{diameter OK})$, which was computed from the full sample space (Table 2.1) to be 0.928.

Example 2.17

Compute the conditional probability $P(\text{diameter OK} \mid \text{length too long})$. Is this the same as the unconditional probability $P(\text{diameter OK})$?

Solution

The conditional probability $P(\text{diameter OK} \mid \text{length too long})$ is computed under the assumption that the rod is too long. This reduces the sample space to the 40 items indicated in boldface in the following table.

	Diameter		
Length	**Too Thin**	**OK**	**Too Thick**
Too Short	10	3	5
OK	38	900	4
Too Long	**2**	**25**	**13**

Of the 40 outcomes, 25 meet the diameter specification. Therefore

$$P(\text{diameter OK} \mid \text{length too long}) = \frac{25}{40} = 0.625$$

The unconditional probability $P(\text{diameter OK})$ is computed on the basis of all 1000 outcomes in the sample space and is equal to $928/1000 = 0.928$. In this case, the conditional probability differs from the unconditional probability.

Let's look at the solution to Example 2.17 more closely. We found that

$$P(\text{diameter OK} \mid \text{length too long}) = \frac{25}{40}$$

In the answer 25/40, the denominator, 40, represents the number of outcomes that satisfy the condition that the rod is too long, while the numerator, 25, represents the number of outcomes that satisfy both the condition that the rod is too long and that its diameter is OK. If we divide both the numerator and denominator of this answer by the number of outcomes in the full sample space, which is 1000, we obtain

$$P(\text{diameter OK} \mid \text{length too long}) = \frac{25/1000}{40/1000}$$

Now 40/1000 represents the *probability* of satisfying the condition that the rod is too long. That is,

$$P(\text{length too long}) = \frac{40}{1000}$$

The quantity 25/1000 represents the *probability* of satisfying both the condition that the rod is too long and that the diameter is OK. That is,

$$P(\text{diameter OK and length too long}) = \frac{25}{1000}$$

We can now express the conditional probability as

$$P(\text{diameter OK} \mid \text{length too long}) = \frac{P(\text{diameter OK and length too long})}{P(\text{length too long})}$$

This reasoning can be extended to construct a definition of conditional probability that holds for any sample space:

Definition

Let A and B be events with $P(B) \neq 0$. The conditional probability of A given B is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{2.14}$$

Figure 2.5 presents Venn diagrams to illustrate the idea of conditional probability.

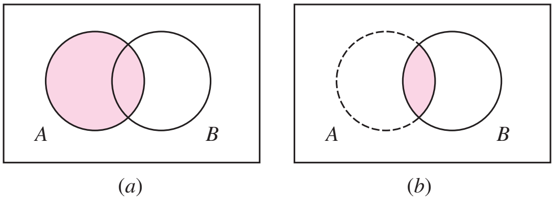


FIGURE 2.5 *(a)* The diagram represents the unconditional probability $P(A)$. $P(A)$ is illustrated by considering the event A in proportion to the entire sample space, which is represented by the rectangle. *(b)* The diagram represents the conditional probability $P(A|B)$. Since the event B is known to occur, the event B now becomes the sample space. For the event A to occur, the outcome must be in the intersection $A \cap B$. The conditional probability $P(A|B)$ is therefore illustrated by considering the intersection $A \cap B$ in proportion to the entire event B.

Example 2.18

Refer to Example 2.8 (in Section 2.1). What is the probability that a can will have a flaw on the side, given that it has a flaw on top?

Solution

We are given that $P(\text{flaw on top}) = 0.03$, and $P(\text{flaw on side and flaw on top}) = 0.01$. Using Equation (2.14),

$$\begin{aligned} P(\text{flaw on side} \mid \text{flaw on top}) &= \frac{P(\text{flaw on side and flaw on top})}{P(\text{flaw on top})} \\ &= \frac{0.01}{0.03} \\ &= 0.33 \end{aligned}$$

Example 2.19

Refer to Example 2.8 (in Section 2.1). What is the probability that a can will have a flaw on the top, given that it has a flaw on the side?

Solution

We are given that $P(\text{flaw on side}) = 0.02$, and $P(\text{flaw on side and flaw on top}) = 0.01$. Using Equation (2.14),

$$\begin{aligned} P(\text{flaw on top} \mid \text{flaw on side}) &= \frac{P(\text{flaw on top and flaw on side})}{P(\text{flaw on side})} \\ &= \frac{0.01}{0.02} \\ &= 0.5 \end{aligned}$$

The results of Examples 2.18 and 2.19 show that in most cases, $P(A|B) \neq P(B|A)$.

Independent Events

Sometimes the knowledge that one event has occurred does not change the probability that another event occurs. In this case the conditional and unconditional probabilities are the same, and the events are said to be **independent**. We present an example.

Example 2.20

If an aluminum rod is sampled from the sample space presented in Table 2.1, find $P(\text{too long})$ and $P(\text{too long} \mid \text{too thin})$. Are these probabilities different?

Solution

$$P(\text{too long}) = \frac{40}{1000} = 0.040$$

$$\begin{aligned} P(\text{too long} \mid \text{too thin}) &= \frac{P(\text{too long and too thin})}{P(\text{too thin})} \\ &= \frac{2/1000}{50/1000} \\ &= 0.040 \end{aligned}$$

The conditional probability and the unconditional probability are the same. The information that the rod is too thin does not change the probability that the rod is too long.

Example 2.20 shows that knowledge that an event occurs sometimes does not change the probability that another event occurs. In these cases, the two events are said to be **independent**. The event that a rod is too long and the event that a rod is too thin are independent. We now give a more precise definition of the term, both in words and in symbols.

Definition

Two events A and B are **independent** if the probability of each event remains the same whether or not the other occurs.

In symbols: If $P(A) \neq 0$ and $P(B) \neq 0$, then A and B are independent if

$$P(B|A) = P(B) \quad \text{or, equivalently,} \quad P(A|B) = P(A) \tag{2.15}$$

If either $P(A) = 0$ or $P(B) = 0$, then A and B are independent.

If A and B are independent, then the following pairs of events are also independent: A and B^c, A^c and B, and A^c and B^c. The proof of this fact is left as an exercise.

The concept of independence can be extended to more than two events:

Definition

Events $A_1, A_2, \ldots, A_n$ are independent if the probability of each remains the same no matter which of the others occur.

In symbols: Events $A_1, A_2, \ldots, A_n$ are independent if for each A_i, and each collection $A_{j1}, \ldots, A_{jm}$ of events with $P(A_{j1} \cap \cdots \cap A_{jm}) \neq 0$,

$$P(A_i|A_{j1} \cap \cdots \cap A_{jm}) = P(A_i) \tag{2.16}$$

The Multiplication Rule

Sometimes we know $P(A|B)$ and we wish to find $P(A \cap B)$. We can obtain a result that is useful for this purpose by multiplying both sides of Equation (2.14) by $P(B)$. This leads to the multiplication rule.

If A and B are two events with $P(B) \neq 0$, then

$$P(A \cap B) = P(B)P(A|B) \tag{2.17}$$

If A and B are two events with $P(A) \neq 0$, then

$$P(A \cap B) = P(A)P(B|A) \tag{2.18}$$

If $P(A) \neq 0$ and $P(B) \neq 0$, then Equations (2.17) and (2.18) both hold.

When two events are independent, then $P(A|B) = P(A)$ and $P(B|A) = P(B)$, so the multiplication rule simplifies:

If A and B are independent events, then

$$P(A \cap B) = P(A)P(B) \tag{2.19}$$

This result can be extended to any number of events. If $A_1, A_2, \ldots, A_n$ are independent events, then for each collection $A_{j1}, \ldots, A_{jm}$ of events

$$P(A_{j1} \cap A_{j2} \cap \cdots \cap A_{jm}) = P(A_{j1})P(A_{j2}) \cdots P(A_{jm}) \tag{2.20}$$

In particular,

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2) \cdots P(A_n) \tag{2.21}$$

Example 2.21

A vehicle contains two engines, a main engine and a backup. The engine component fails only if both engines fail. The probability that the main engine fails is 0.05, and the probability that the backup engine fails is 0.10. Assume that the main and backup engines function independently. What is the probability that the engine component fails?

Solution

The probability that the engine component fails is the probability that both engines fail. Therefore

$$P(\text{engine component fails}) = P(\text{main engine fails and backup engine fails})$$

Since the engines function independently, we may use Equation (2.19):

$$\begin{aligned} P(\text{main engine fails and backup engine fails}) &= P(\text{main fails})P(\text{backup fails}) \\ &= (0.10)(0.05) \\ &= 0.005 \end{aligned}$$

Example 2.22

A system contains two components, A and B. Both components must function for the system to work. The probability that component A fails is 0.08, and the probability that component B fails is 0.05. Assume the two components function independently. What is the probability that the system functions?

Solution

The probability that the system functions is the probability that both components function. Therefore

$$P(\text{system functions}) = P(\text{A functions and B functions})$$

Since the components function independently,

$$\begin{aligned} P(\text{A functions and B functions}) &= P(\text{A functions})P(\text{B functions}) \\ &= [1 - P(\text{A fails})][1 - P(\text{B fails})] \\ &= (1 - 0.08)(1 - 0.05) \\ &= 0.874 \end{aligned}$$

Example 2.23

Of the microprocessors manufactured by a certain process, 20% are defective. Five microprocessors are chosen at random. Assume they function independently. What is the probability that they all work?

Solution

For $i = 1, \ldots, 5$, let A_i denote the event that the ith microprocessor works. Then

$$\begin{aligned} P(\text{all 5 work}) &= P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) \\ &= P(A_1)P(A_2)P(A_3)P(A_4)P(A_5) \\ &= (1 - 0.20)^5 \\ &= 0.328 \end{aligned}$$

Example 2.24

In Example 2.23, what is the probability that at least one of the microprocessors works?

Solution

The easiest way to solve this problem is to notice that

$$P(\text{at least one works}) = 1 - P(\text{all are defective})$$

Now, letting D_i denote the event that the ith microprocessor is defective,

$$\begin{aligned}P(\text{all are defective}) &= P(D_1 \cap D_2 \cap D_3 \cap D_4 \cap D_5) \\ &= P(D_1)P(D_2)P(D_3)P(D_4)P(D_5) \\ &= (0.20)^5 \\ &= 0.0003\end{aligned}$$

Therefore $P(\text{at least one works}) = 1 - 0.0003 = 0.9997$.

Equations (2.19) and (2.20) tell us how to compute probabilities when we know that events are independent, but they are usually not much help when it comes to deciding whether two events really *are* independent. In most cases, the best way to determine whether events are independent is through an understanding of the process that produces the events. Here are a few illustrations:

- A die is rolled twice. It is reasonable to believe that the outcome of the second roll is not affected by the outcome of the first roll. Therefore, knowing the outcome of the first roll does not help to predict the outcome of the second roll. The two rolls are independent.
- A certain chemical reaction is run twice, using different equipment each time. It is reasonable to believe that the yield of one reaction will not affect the yield of the other. In this case the yields are independent.
- A chemical reaction is run twice in succession, using the same equipment. In this case, it might not be wise to assume that the yields are independent. For example, a low yield on the first run might indicate that there is more residue than usual left behind. This might tend to make the yield on the next run higher. Thus knowing the yield on the first run could help to predict the yield on the second run.
- The items in a simple random sample may be treated as independent, unless the population is finite and the sample comprises more than about 5% of the population (see the discussion of independence in Section 1.1).

The Law of Total Probability

The law of total probability is illustrated in Figure 2.6. A sample space contains the events A_1, A_2, A_3, and A_4. These events are mutually exclusive, since no two overlap. They are also **exhaustive**, which means that their union covers the whole sample space. Each outcome in the sample space belongs to one and only one of the events A_1, A_2, A_3, A_4.

The event B can be any event. In Figure 2.6, each of the events A_i intersects B, forming the events $A_1 \cap B$, $A_2 \cap B$, $A_3 \cap B$, and $A_4 \cap B$. It is clear from Figure 2.6 that the events $A_1 \cap B$, $A_2 \cap B$, $A_3 \cap B$, and $A_4 \cap B$ are mutually exclusive and that they cover B. Every outcome in B belongs to one and only one of the events $A_1 \cap B$, $A_2 \cap B$, $A_3 \cap B$, $A_4 \cap B$. It follows that

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup (A_3 \cap B) \cup (A_4 \cap B)$$