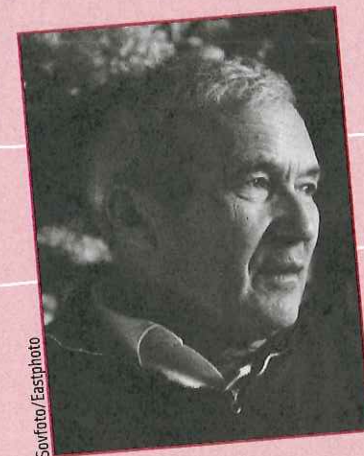
PART III

Probability:
Foundations for Inference



Sovfoto/Eastphoto

A. N. KOLMOGOROV

General Laws of Probability

There are national styles in science as well as in cuisine. Statistics, the science of data, was created mainly by British and Americans. Probability, the mathematics of chance, was long led by French and Russians. *Andrei Nikolaevich Kolmogorov (1903–1987)* was the greatest of the Russian probabilists and one of the most influential mathematicians of the twentieth century. His more than 500 mathematical publications shaped several areas of modern mathematics and applied mathematical ideas to areas as far afield as the rhythms and meters of poetry.

Kolmogorov entered Moscow State University as a student in 1920 and remained there until his death. He was named a Hero of Socialist Labor in 1963, a rare honor for someone whose career was devoted entirely to scholarship.

Kolmogorov's first work in probability concerned the behavior of strings of random observations. The law of large numbers is the starting point for these studies, and Kolmogorov discovered many extensions of that law. Kolmogorov effectively established probability as a field of mathematics in 1933, when he placed it on a firm mathematical foundation by starting with a few general laws from which all else follows. The general laws of probability in this chapter are in the spirit of Kolmogorov.

Statistics, the science of data, was created mainly by British and Americans. Probability, the mathematics of chance, was long led by French and Russians.

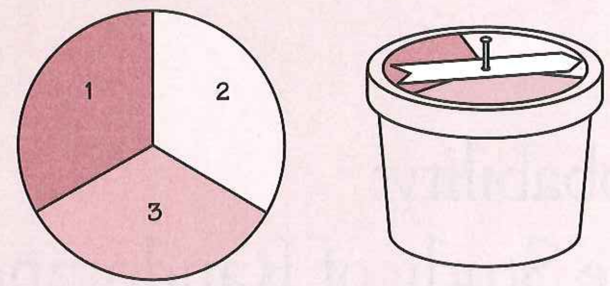
chapter 6

Probability: The Study of Randomness

ACTIVITY 6 The Spinning Wheel

Materials: Margarine tub spinner or graphing calculator or table of random numbers

Imagine a spinner with three sectors, all the same size, marked 1, 2, and 3 as shown.



The experiment consists of spinning the spinner three times and recording the numbers as they occur (e.g., 123). We want to determine the proportion of times that *at least one digit occurs in its correct position*. For example, in the number 123, all of the digits are in their proper positions, but in the number 331, none are. For this activity, use a spinner like the one in the illustration, a table of random digits, or your calculator.

1. Guess the proportion of times at least one digit will occur in its proper place.

2. To use your calculator to randomly generate the three-digit number, enter the command `randInt(1,3,3)`. Continue to press ENTER to generate more three-digit numbers. Use a tally mark to record the results in a table like the one below. Do 20 trials and then calculate the relative frequency for the event "at least one digit in the correct position."

At least one digit in the correct position	
Not	

To use a random number table, select a row, and discarding digits 4 to 9 and 0, record digits in the 1 to 3 range in groups of three.

ACTIVITY 6 The Spinning Wheel (*continued*)

3. Combine your results with those of your classmates to obtain as many trials as possible (at least 100 randomly generated three-digit numbers; 200 would be better).

4. Count the number of times at least one digit occurred in its correct position, and calculate the proportion.

5. The program SPIN123 implements the experiment for the TI-83/89. The key step uses the calculator's Boolean logic to count the number of "hits." Enter the program or link it from a classmate or your teacher.

TI-83

```
PROGRAM:SPIN123
:ClrHome
:ClrList L1,L2
:Disp "HOW MANY TRIALS"
:Prompt N
:1→C
:While C≤N
:randInt(1,3,3)→L1
:(L1(1)=1 or L1(2)=2 or
 L1(3)=3)→L2(C)
:1+C→C
:End
:Disp "REL FREQ="
:Disp sum(L2=1)/N
```

TI-89

```
spin123()
Prgm
ClrHome
tistat.clrlist(list1,
list2)
Disp "how many trials"
Prompt n
1→c
While c≤n
tistat.randint(1,3,3)→
 list1
list1[1]=1 or list1[2]=2
 or list1[3]=3→list2[c]
1+c→c
EndWhile
Disp "rel freq="
0→s
For i,1,n
If list2[i]=true
s+1→s
EndFor
Disp s/n
```

Execute the program for 25, 50, and 100 repetitions. Compare the calculator results with the results you obtained in steps 2 to 4.

Later in the chapter we will calculate the theoretical probability of this event happening, so keep your data at hand so that you can compare the theoretical probability with your experimental results.

INTRODUCTION

Chance is all around us. Sometimes chance results from human design, as in the casino's games of chance and the statistician's random samples. Sometimes nature uses chance, as in choosing the sex of a child. Sometimes the reasons for chance behavior are mysterious, as when the number of deaths each year in a large population is as regular as the number of heads in many tosses of a coin. Probability is the branch of mathematics that describes the pattern of chance outcomes.

The reasoning of statistical inference rests on asking, "How often would this method give a correct answer if I used it very many times?" When we produce data by random sampling or randomized comparative experiments, the laws of probability answer the question "What would happen if we did this many times?" This chapter presents the fundamental concepts of probability. Probability calculations are the basis for inference. The tools you acquire in this chapter will help you describe the behavior of statistics from random samples and randomized comparative experiments in later chapters. Even our brief acquaintance with probability will enable us to answer questions like these:

- If we know the blood types of a man and a woman, what can we say about the blood types of their future children?
- Give a test for the AIDS virus to the employees of a small company. What is the chance of at least one positive test if all the people tested are free of the virus?
- An opinion poll asks a sample of 1500 adults what they consider the most serious problem facing our schools. How often will the poll percent who answer "drugs" come within two percentage points of the truth about the entire population?

6.1 THE IDEA OF PROBABILITY

The mathematics of probability begins with the observed fact that some phenomena are random—that is, the relative frequencies of their outcomes seem to settle down to fixed values in the long run. Consider tossing a single coin. The relative frequency of heads is quite erratic in 2 or 5 or 10 tosses. But after several thousand tosses it remains stable, changing very little over further thousands of tosses. The big idea is this: **chance behavior is unpredictable in the short run but has a regular and predictable pattern in the long run.**

Toss a coin, or choose an SRS. The result can't be predicted in advance, because the result will vary when you toss the coin or choose the sample repeatedly. But there is still a regular pattern in the results, a pattern that emerges clearly only after many repetitions. This remarkable fact is the basis for the idea of probability.

EXAMPLE 6.1 COIN TOSSING

When you toss a coin, there are only two possible outcomes, heads or tails. Figure 6.1 shows the results of tossing a coin 1000 times. For each number of tosses from 1 to 1000, we have plotted the proportion of those tosses that gave a head. The first toss was a head, so the proportion of heads starts at 1. The second toss was a tail, reducing the proportion of heads to 0.5 after two tosses. The next three tosses gave a tail followed by two heads, so the proportion of heads after five tosses is 3/5, or 0.6.

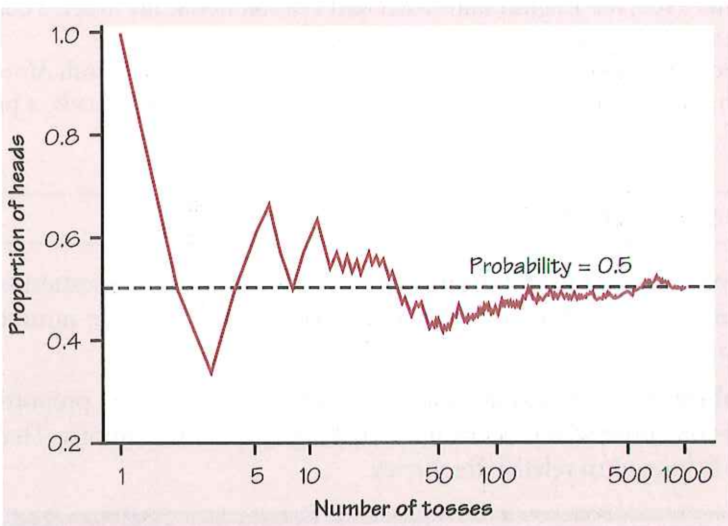


FIGURE 6.1 The behavior of the proportion of coin tosses that give a head, from 1 to 1000 tosses of a coin. In the long run, the proportion of heads approaches 0.5, the probability of a head.

The proportion of tosses that produce heads is quite variable at first, but it settles down as we make more and more tosses. Eventually this proportion gets close to 0.5 and stays there. We say that 0.5 is the *probability* of a head. The probability 0.5 appears as a horizontal line on the graph.

"Random" in statistics is not a synonym for "haphazard" but a description of a kind of order that emerges only in the long run. We often encounter the unpredictable side of randomness in our everyday experience, but we rarely see enough repetitions of the same random phenomenon to observe the long-term regularity that probability describes. You can see that regularity emerging in Figure 6.1. In the very long run, the proportion of tosses that give a head is 0.5. This is the intuitive idea of probability. Probability 0.5 means "occurs half the time in a very large number of trials."

We might suspect that a coin has probability 0.5 of coming up heads just because the coin has two sides. As Exercise 6.1 illustrates, such suspicions are not always correct. The idea of probability is empirical. That is, it is based on observation rather than theorizing. Probability describes what happens in very many trials, and we must actually observe many trials to pin down a probability. In the case of tossing a coin, some diligent people have in fact made thousands of tosses.

EXAMPLE 6.2 SOME COIN TOSSERS

The French naturalist Count Buffon (1707–1788) tossed a coin 4040 times. Result: 2048 heads, or proportion 2048/4040 = 0.5069 for heads.

Around 1900, the English statistician Karl Pearson heroically tossed a coin 24,000 times. Result: 12,012 heads, a proportion of 0.5005.

While imprisoned by the Germans during World War II, the South African mathematician John Kerrich tossed a coin 10,000 times. Result: 5067 heads, a proportion of 0.5067.

RANDOMNESS AND PROBABILITY

We call a phenomenon **random** if individual outcomes are uncertain but there is nonetheless a regular distribution of outcomes in a large number of repetitions.

The **probability** of any outcome of a random phenomenon is the proportion of times the outcome would occur in a very long series of repetitions. That is, probability is long-term relative frequency.

Thinking about randomness

That some things are random is an observed fact about the world. The outcome of a coin toss, the time between emissions of particles by a radioactive source, and the sexes of the next litter of lab rats are all random. So is the outcome of a random sample or a randomized experiment. Probability theory is the branch of mathematics that describes random behavior. Of course, we can never observe a probability exactly. We could always continue tossing the coin, for example. Mathematical probability is an idealization based on imagining what would happen in an indefinitely long series of trials.

The best way to understand randomness is to observe random behavior—not only the long-run regularity but the unpredictable results of short runs. You can do this with physical devices, as in Exercises 6.1, 6.2, 6.6, and 6.7, but computer simulations (imitations) of random behavior allow faster exploration. Exercises 6.3 and 6.10 suggest some simulations of random behavior. As you explore randomness, remember:

independence

- You must have a long series of ***independent*** trials. That is, the outcome of one trial must not influence the outcome of any other. Imagine a crooked gambling house where the operator of a roulette wheel can stop it where she chooses—she can prevent the proportion of "red" from settling down to a fixed number. These trials are not independent.

- The idea of probability is empirical. Computer simulations start with given probabilities and imitate random behavior, but we can estimate a real-world probability only by actually observing many trials.

- Nonetheless, computer simulations are very useful because we need long runs of trials. In situations such as coin tossing, the proportion of an outcome often requires several hundred trials to settle down to the probability of that outcome. The kinds of physical random devices suggested in the exercises are too slow for this. Short runs give only rough estimates of a probability.

The uses of probability

Probability theory originated in the study of games of chance. Tossing dice, dealing shuffled cards, and spinning a roulette wheel are examples of deliberate randomization that are similar to random sampling. Although games of chance are ancient, they were not studied by mathematicians until the sixteenth and seventeenth centuries. It is only a mild simplification to say that probability as a branch of mathematics arose when seventeenth-century French gamblers asked the mathematicians Blaise Pascal and Pierre de Fermat for help. Gambling is still with us, in casinos and state lotteries. We will make use of games of chance as simple examples that illustrate the principles of probability.

Careful measurements in astronomy and surveying led to further advances in probability in the eighteenth and nineteenth centuries because the results of repeated measurements are random and can be described by distributions much like those arising from random sampling. Similar distributions appear in data on human life span (mortality tables) and in data on lengths or weights in a population of skulls, leaves, or cockroaches.[1] In the twentieth century, we employ the mathematics of probability to describe the flow of traffic through a highway system, a telephone interchange, or a computer processor; the genetic makeup of individuals or populations; the energy states of subatomic particles; the spread of epidemics or rumors; and the rate of return on risky investments. Although we are interested in probability because of its usefulness in statistics, the mathematics of chance is important in many fields of study.

SECTION 6.1 EXERCISES

6.1 **PENNIES SPINNING** Hold a penny upright on its edge under your forefinger on a hard surface, then snap it with your other forefinger so that it spins for some time before falling. Based on 50 spins, estimate the probability of heads.

6.2 A GAME OF CHANCE In the game of Heads or Tails, Betty and Bob toss a coin four times. Betty wins a dollar from Bob for each head and pays Bob a dollar for each tail—that is, she wins or loses the difference between the number of heads and the number of tails. For example, if there are one head and three tails, Betty loses \$2. You can check that Betty's possible outcomes are

$$\{-4, -2, 0, 2, 4\}$$

Assign probabilities to these outcomes by playing the game 20 times and using the proportions of the outcomes as estimates of the probabilities. If possible, combine your trials with those of other students to obtain long-run proportions that are closer to the probabilities.

6.3 SHAQ The basketball player Shaquille O'Neal makes about half of his free throws over an entire season. We will use the calculator to simulate 100 free throws shot independently by a player who has probability 0.5 of making each shot. We let the number 1 represent the outcome "Hit" and 0 represent a "Miss."

(a) Enter the command `randInt(0,1,100)→SHAQ`. (`randInt` is found in the CATALOG under Flash Apps on the TI-89.) This tells the calculator to randomly select a hit (1) or a miss (0), do this 100 times in succession, and store the results in the list named SHAQ.

(b) What percent of the 100 shots are hits?

(c) Examine the sequence of hits and misses. How long was the longest run of shots made? Of shots missed? (Sequences of random outcomes often show runs longer than our intuition thinks likely.)

6.4 MATCHING PROBABILITIES Probability is a measure of how likely an event is to occur. Match one of the probabilities that follow with each statement about an event. (The probability is usually a much more exact measure of likelihood than is the verbal statement.)

$$0, 0.01, 0.3, 0.6, 0.99, 1$$

(a) This event is impossible. It can never occur.

(b) This event is certain. It will occur on every trial of the random phenomenon.

(c) This event is very unlikely, but it will occur once in a while in a long sequence of trials.

(d) This event will occur more often than not.

6.5 RANDOM DIGITS The table of random digits (Table B) was produced by a random mechanism that gives each digit probability 0.1 of being a 0. What proportion of the first 200 digits in the table are 0s? This proportion is an estimate, based on 200 repetitions, of the true probability, which in this case is known to be 0.1.

6.6 HOW MANY TOSSES TO GET A HEAD? When we toss a penny, experience shows that the probability (long-term proportion) of a head is close to 1/2. Suppose now that we toss the penny repeatedly until we get a head. What is the probability that the first head comes up in an odd number of tosses (1, 3, 5, and so on)? To find out, repeat this experiment 50 times, and keep a record of the number of tosses needed to get a head on each of your 50 trials.

(a) From your experiment, estimate the probability of a head on the first toss. What value should we expect this probability to have?

(b) Use your results to estimate the probability that the first head appears on an odd-numbered toss.

6.7 **TOSSING A THUMBTACK** Toss a thumbtack on a hard surface 100 times. How many times did it land with the point up? What is the approximate probability of landing point up?

6.8 **THREE OF A KIND** You read in a book on poker that the probability of being dealt three of a kind in a five-card poker hand is 1/50. Explain in simple language what this means.

6.9 **WINNING A BASEBALL GAME** A study of the home-field advantage in baseball found that over the period from 1969 to 1989 the league champions won 63% of their home games.[2] The two league champions meet in the baseball World Series. Would you use the study results to assign probability 0.63 to the event that the home team wins in a World Series game? Explain your answer.

6.10 **SIMULATING AN OPINION POLL** A recent opinion poll showed that about 73% of married women agree that their husbands do at least their fair share of household chores. Suppose that this is exactly true. Choosing a married woman at random then has probability 0.73 of getting one who agrees that her husband does his share. Use software or your calculator to simulate choosing many women independently. (In most software, the key phrase to look for is "Bernoulli trials." This is the technical term for independent trials with Yes/No outcomes. Our outcomes here are "Agree" or not.)

(a) Simulate drawing 20 women, then 80 women, then 320 women. What proportion agree in each case? We expect (but because of chance variation we can't be sure) that the proportion will be closer to 0.73 in longer runs of trials.

(b) Simulate drawing 20 women 10 times and record the percents in each trial who agree. Then simulate drawing 320 women 10 times and again record the 10 percents. Which set of 10 results is less variable? We expect the results of 320 trials to be more predictable (less variable) than the results of 20 trials. That is "long-run regularity" showing itself.

6.2 PROBABILITY MODELS

Earlier chapters gave mathematical models for linear relationships (in the form of the equation of a line) and for some distributions of data (in the form of normal density curves). Now we must give a mathematical description or model for randomness. To see how to proceed, think first about a very simple random phenomenon, tossing a coin once. When we toss a coin, we cannot know the outcome in advance. What do we know? We are willing to say that the outcome will be either heads or tails. We believe that each of these outcomes has probability 1/2. This description of coin tossing has two parts:

- A list of possible outcomes.
- A probability for each outcome.

Such a description is the basis for all probability models. Here is the basic vocabulary we use.

PROBABILITY MODELS

The **sample space** S of a random phenomenon is the set of all possible outcomes.

An **event** is any outcome or a set of outcomes of a random phenomenon. That is, an event is a subset of the sample space.

A **probability model** is a mathematical description of a random phenomenon consisting of two parts: a sample space S and a way of assigning probabilities to events.

The sample space S can be very simple or very complex. When we toss a coin once, there are only two outcomes, heads and tails. The sample space is $S = \{H, T\}$. If we draw a random sample of 50,000 U.S. households, as the Current Population Survey does, the sample space contains all possible choices of 50,000 of the 103 million households in the country. This S is extremely large. Each member of S is a possible sample, which explains the term *sample space*.

EXAMPLE 6.3 ROLLING DICE

Rolling two dice is a common way to lose money in casinos. There are 36 possible outcomes when we roll two dice and record the up-faces in order (first die, second die). Figure 6.2 displays these outcomes. They make up the sample space S.

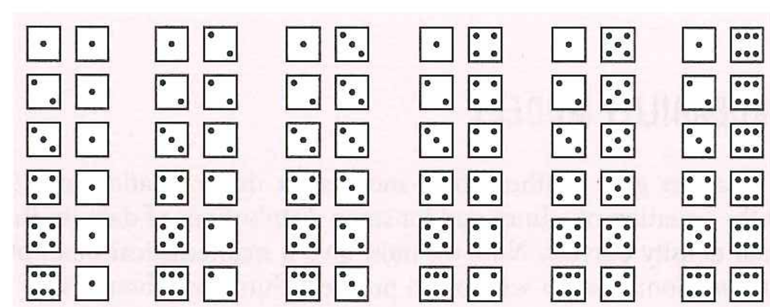


FIGURE 6.2 The 36 possible outcomes in rolling two dice.

"Roll a 5" is an event, call it A, that contains four of these 36 outcomes:

A = {  }

Gamblers care only about the number of pips on the up-faces of the dice. The sample space for rolling two dice and counting the pips is

$$S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

Comparing this S with Figure 6.2 reminds us that we can change S by changing the detailed description of the random phenomenon we are describing.

The name "sample space" is natural in random sampling, where each possible outcome is a sample and the sample space contains all possible samples.

To specify S, we must state what constitutes an individual outcome and then state which outcomes can occur. We often have some freedom in defining the sample space, so the choice of S is a matter of convenience as well as correctness. The idea of a sample space, and the freedom we may have in specifying it, are best illustrated by examples.

EXAMPLE 6.4 RANDOM DIGIT

Let your pencil point fall blindly into Table B of random digits; record the value of the digit it lands on. The possible outcomes are

$$S = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

EXAMPLE 6.5 FLIP A COIN AND ROLL A DIE

An experiment consists of flipping a coin and rolling a die. Possible outcomes are a head (H) followed by any of the digits 1 to 6, or a tail (T) followed by any of the digits 1 to 6. The sample space contains 12 outcomes:

$$S = \{H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6\}$$

tree diagram

Being able to properly enumerate the outcomes in a sample space will be critical to determining probabilities. Two techniques are very helpful in making sure you don't accidentally overlook any outcomes. The first is called a ***tree diagram*** because it resembles the branches of a tree. The first action in Example 6.5 is to toss a coin. To construct the tree diagram, begin with a point and draw a line from the point to H and a second line from the point to T. The second action is to roll a die; there are six possible faces that can come up on the die. So draw a line from each of H and T to these six outcomes. See Figure 6.3.

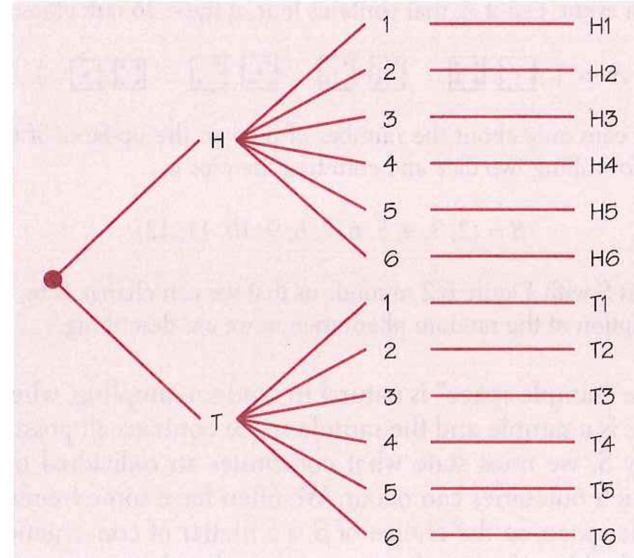


FIGURE 6.3 Tree diagram.

The second technique is to make use of the following rule.

MULTIPLICATION PRINCIPLE

If you can do one task in a number of ways and a second task in b number of ways, then both tasks can be done in $a \times b$ number of ways.

To determine the number of outcomes in the sample space for Example 6.5, there are 2 ways the coin can come up, and there are 6 ways the die can come up, so there are 2×6 possible outcomes in the sample space. To see why this is true, just sketch a tree diagram.

EXAMPLE 6.6 FLIP FOUR COINS

An experiment consists of flipping four coins. You can think of either tossing four coins onto the table all at once or flipping a coin four times in succession and recording the four outcomes. One possible outcome is HHTH. Because there are two ways each coin can come up, the multiplication principle says that the total number of outcomes is $2 \times 2 \times 2 \times 2 = 16$. This is the easy part. Listing all 16 outcomes requires a scheme or systematic method so that you don't leave out any possibilities. One way is to list all the ways you can obtain 0 heads, then list all the ways you can get 1 head, 2 heads, 3 heads, and finally all 4 heads. Here is an enumeration:

0 heads	1 head	2 heads	3 heads	4 heads
TTTT	HTTT	HHTT	HHHT	HHHH
	THTT	HTHT	HHTH	
	TTHT	HTTH	HTHH	
	TTTH	THHT	THHH	
		THTH		
		TTHH		

Suppose that our only interest is the number of heads in four tosses. Now we can be exact in a simpler fashion. The random phenomenon is to toss a coin four times and count the number of heads. The sample space contains only five outcomes:

$$S = \{0, 1, 2, 3, 4\}$$

This example also illustrates the importance of carefully specifying what constitutes an individual outcome.

Although these examples seem remote from the practice of statistics, the connection is surprisingly close. Suppose that in the course of conducting an opinion poll you select four people at random from a large population and ask each if he or she favors reducing federal spending on low-interest student loans. The possible outcomes—the sample space—are the answers "Yes" or "No." Similarly, the possible outcomes of an SRS of 1500 people are the same in principle as the possible outcomes of tossing a coin 1500 times. One of the great advantages of mathematics is that the essential features of quite different phenomena can be described by the same mathematical model.

Of course, some sample spaces are simply too large to allow all of the possible outcomes to be listed, as the next example shows.

EXAMPLE 6.7 GENERATE A RANDOM DECIMAL NUMBER

Many computing systems have a function that will generate a random number between 0 and 1. The sample space is

$$S = \{\text{all numbers between 0 and 1}\}$$

This S is a mathematical idealization. Any specific random number generator produces numbers with some limited number of decimal places so that, strictly speaking, not all numbers between 0 and 1 are possible outcomes. The entire interval from 0 to 1 is easier to think about. It also has the advantage of being a suitable sample space for different computers that produce random numbers with different numbers of significant digits.

If you are selecting objects from a collection of distinct choices, such as drawing playing cards from a standard deck of 52 cards, then much depends on whether each choice is exactly like the previous choice. If you are selecting random digits by drawing numbered slips of paper from a hat, and you want all ten digits to be equally likely to be selected each draw, then after you draw a digit and record it, you must put it back into the hat. Then the second draw will be exactly like the first. This is referred to as sampling ***with replacement.*** If you do not replace the slips you draw, however, there are only nine choices for the second slip picked, and eight for the third. This is called sampling ***without replacement***. So if the question is "How many three-digit numbers can you make?" the answer is, by the multiplication principle, $10 \times 10 \times 10 = 1000$, providing all ten numbers are eligible for each of the three positions in the number. On the other had, there are $10 \times 9 \times 8 = 720$ different ways to construct a three-digit number *without replacement*. You should be able to determine from the context of the problem whether the selection is with or without replacement, and this will help you properly identify the sample space.

replacement

EXERCISES

6.11 DESCRIBE THE SAMPLE SPACE In each of the following situations, describe a sample space S for the random phenomenon. In some cases, you have some freedom in your choice of S.

(a) A seed is planted in the ground. It either germinates or fails to grow.

(b) A patient with a usually fatal form of cancer is given a new treatment. The response variable is the length of time that the patient lives after treatment.

(c) A student enrolls in a statistics course and at the end of the semester receives a letter grade.

(d) A basketball player shoots four free throws. You record the sequence of hits and misses.

(e) A basketball player shoots four free throws. You record the number of baskets she makes.

6.12 DESCRIBE THE SAMPLE SPACE In each of the following situations, describe a sample space S for the random phenomenon. In some cases you have some freedom in specifying S, especially in setting the largest and the smallest value in S.

(a) Choose a student in your class at random. Ask how much time that student spent studying during the past 24 hours.

(b) The Physicians' Health Study asked 11,000 physicians to take an aspirin every other day and observed how many of them had a heart attack in a five-year period.

(c) In a test of a new package design, you drop a carton of a dozen eggs from a height of 1 foot and count the number of broken eggs.

(d) Choose a student in your class at random. Ask how much cash that student is carrying.

(e) A nutrition researcher feeds a new diet to a young male white rat. The response variable is the weight (in grams) that the rat gains in 8 weeks.

6.13 **CALORIES IN HOT DOGS** Give a reasonable sample space for the number of calories in a hot dog. (Table 1.10 on page 59 contains some typical values to guide you.)

6.14 **LISTING OUTCOMES, I** For each of the following, use a tree diagram or the multiplication principle to determine the number of outcomes in the sample space. Then write the sample space using set notation.

(a) Toss 2 coins.

(b) Toss 3 coins.

(c) Toss 4 coins.

6.15 **LISTING OUTCOMES, II** For each of the following, use a tree diagram or the multiplication principle to determine the number of outcomes in the sample space.

(a) Suppose a county license tag has a four-digit number for identification. If any digit can occupy any of the four positions, how many county license tags can you have?

(b) If the county license tags described in (a) do not allow duplicate digits, how many county license tags can you have?

(c) Suppose the county license tags described in (a) can have *up* to four digits. How many county license tags will this scheme allow?

6.16 **SPIN 123** Refer to the experiment described in Activity 6.

(a) Determine the number of outcomes in the sample space.

(b) List the outcomes in the sample space.

6.17 **ROLLING TWO DICE** Example 6.3 (page 336) showed the 36 outcomes when we roll two dice. Another way to summarize these results is to make a table like this:

Number of ways	Sum	Outcomes
1	2	1,1
2	3	1,2 2,1
...	...	...

(a) Complete the table.

(b) In how many ways can you get an even sum?

(c) In how many ways can you get a sum of 5? A sum of 8?

(d) Describe any patterns that you see in the table.

6.18 PICK A CARD Suppose you select a card from a standard deck of 52 playing cards. In how many ways can the selected card be

(a) a red card?

(b) a heart?

(c) a queen and a heart?

(d) a queen or a heart?

(e) a queen that is not a heart?

Probability rules

The true probability of any outcome—say, "roll a 5 when we toss two dice"—can be found only by actually tossing two dice many times, and then only approximately. How then can we describe probability mathematically? Rather than try to give "correct" probabilities, we start by laying down facts that must be true for any assignment of probabilities. These facts follow from the idea of probability as "the long-run proportion of repetitions on which an event occurs."

1. Any probability is a number between 0 and 1. Any proportion is a number between 0 and 1, so any probability is also a number between 0 and 1. An event with probability 0 never occurs, and an event with probability 1 occurs on every trial. An event with probability 0.5 occurs in half the trials in the long run.

2. All possible outcomes together must have probability 1. Because some outcome must occur on every trial, the sum of the probabilities for all possible outcomes must be exactly 1.

3. The probability that an event does not occur is 1 minus the probability that the event does occur. If an event occurs in (say) 70% of all trials, it fails to occur in the other 30%. The probability that an event occurs and the probability that it does not occur always add to 100%, or 1.

4. If two events have no outcomes in common, the probability that one or the other occurs is the sum of their individual probabilities. If one event occurs in 40% of all trials, a different event occurs in 25% of all trials, and the two can never occur together, then one or the other occurs on 65% of all trials because 40% + 25% = 65%.

We can use mathematical notation to state Facts 1 to 4 more concisely. Capital letters near the beginning of the alphabet denote events. If A is any event, we write its probability as $P(A)$. Here are our probability facts in formal language. As you apply these rules, remember that they are just another form of intuitively true facts about long-run proportions.

PROBABILITY RULES

Rule 1. The probability $P(A)$ of any event A satisfies $0 \leq P(A) \leq 1$.

Rule 2. If S is the sample space in a probability model, then $P(S) = 1$.

Rule 3. The **complement** of any event A is the event that A does not occur, written as A^c. The **complement rule** states that

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

Rule 4. Two events A and B are **disjoint** (also called mutually exclusive) if they have no outcomes in common and so can never occur simultaneously. If A and B are disjoint,

$$P(A \text{ or } B) = P(A) + P(B)$$

This is the **addition rule** for disjoint events.

Sometime we use set notation to describe events. The event $\{A \cup B\}$, read "A ***union*** B," is the set of all outcomes that are either in A or in B. So $\{A \cup B\}$ is just another way to indicate the event {A or B}. We will use these two notations interchangeably. The symbol $\varnothing$ is used for the ***empty event***, that is, the event that has no outcomes in it. If two events A and B are disjoint (mutually exclusive), we can write $A \cap B = \varnothing$, read "A ***intersect*** B is empty." Sometimes we emphasize that we are describing a compound event by enclosing it within braces.

union

empty event

intersect

You may find it helpful to draw a picture to remind yourself of the meaning of complements and disjoint events. A picture like Figure 6.4 that shows the sample space S as a rectangular area and events as areas within S is called a ***Venn diagram***. The events A and B in Figure 6.4 are disjoint because they do not overlap; that is, they have no outcomes in common. Their intersection is the empty event, $\varnothing$. Their union consists of the two shaded regions.

Venn diagram

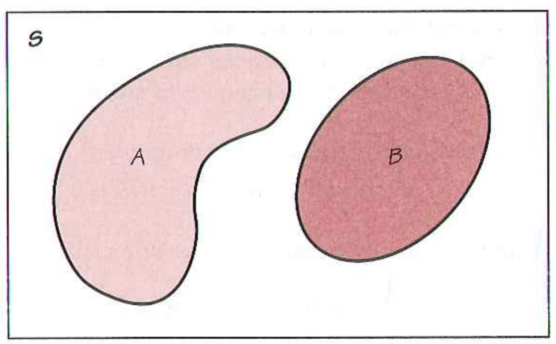


FIGURE 6.4 Venn diagram showing disjoint (mutually exclusive) events *A* and *B*.

The complement A^c in Figure 6.5 contains exactly the outcomes that are not in A. Note that we could write $A \cup A^c = S$ and $A \cap A^c = \emptyset$.

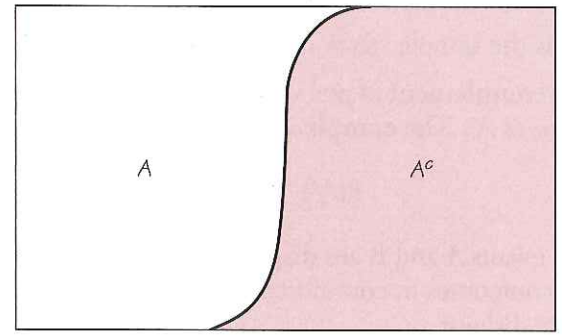


FIGURE 6.5 Venn diagram showing the complement A^c of an event A.

EXAMPLE 6.8 MARITAL STATUS OF YOUNG WOMEN

Draw a woman aged 25 to 34 years old at random and record her marital status. "At random" means that we give every such woman the same chance to be the one we choose. That is, we choose an SRS of size 1. The probability of any marital status is just the proportion of all women aged 25 to 34 who have that status—if we drew many women, this is the proportion we would get. Here is the probability model:

Marital status:	Never married	Married	Widowed	Divorced
Probability:	0.298	0.622	0.005	0.075

Each probability is between 0 and 1. The probabilities add to 1 because these outcomes together make up the sample space S.

The probability that the woman we draw is not married is, by the complement rule,

$$\begin{aligned} P(\text{not married}) &= 1 - P(\text{married}) \\ &= 1 - 0.622 = 0.378 \end{aligned}$$

That is, if 62.2% are married, then the remaining 37.8% are not married.

"Never married" and "divorced" are disjoint events, because no woman can be both never married and divorced. So the addition rule says that

$$\begin{aligned} P(\text{never married or divorced}) &= P(\text{never married}) + P(\text{divorced}) \\ &= 0.298 + 0.075 = 0.373 \end{aligned}$$

That is, 37.3% of women in this age group are either never married or divorced.

EXAMPLE 6.9 PROBABILITIES FOR ROLLING DICE

Figure 6.2 (page 336) displays the 36 possible outcomes of rolling two dice. What probabilities should we assign to these outcomes?

Casino dice are carefully made. Their spots are not hollowed out, which would give the faces different weights, but are filled with white plastic of the same density as the colored plastic of the body. For casino dice it is reasonable to assign the same probability to each of the 36 outcomes in Figure 6.2. Because all 36 outcomes together must have probability 1 (Rule 2), each outcome must have probability 1/36.

Gamblers are often interested in the sum of the pips on the up-faces. What is the probability of rolling a 5? Because the event "roll a 5" contains the four outcomes displayed in Example 6.3, the addition rule (Rule 4) says that its probability is

$$P(\text{roll a } 5) = P(\text{⚀⚃}) + P(\text{⚂⚁}) + P(\text{⚁⚂}) + P(\text{⚃⚀})$$

$$= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36}$$

$$= \frac{4}{36} = 0.111$$

What about the probability of rolling a 7? In Figure 6.2 you will find six outcomes for which the sum of the pips is 7. The probability is 6/36, or about 0.167.

Assigning probabilities: finite number of outcomes

Examples 6.8 and 6.9 illustrate one way to assign probabilities to events: assign a probability to every individual outcome, then add these probabilities to find the probability of any event. If such an assignment is to satisfy the rules of probability, the probabilities of all the individual outcomes must sum to exactly 1.

PROBABILITIES IN A FINITE SAMPLE SPACE

Assign a probability to each individual outcome. These probabilities must be numbers between 0 and 1 and must have sum 1.

The probability of any event is the sum of the probabilities of the outcomes making up the event.

EXAMPLE 6.10 BENFORD'S LAW

Faked numbers in tax returns, payment records, invoices, expense account claims, and many other settings often display patterns that aren't present in legitimate records. Some patterns, like too many round numbers, are obvious and easily avoided by a clever crook. Others are more subtle. It is a striking fact that the first digits of numbers in legitimate records often follow a distribution known as *Benford's Law*. Here it is (note that a first digit can't be 0):[3]

First digit:	1	2	3	4	5	6	7	8	9
Probability:	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046

Benford's Law usually applies to the first digits of the sizes of similar quantities, such as invoices, expense account claims, and county populations. Investigators can detect fraud by comparing these probabilities with the first digits in records such as invoices paid by a business.

Consider the events

$$A = \{\text{first digit is } 1\}$$
$$B = \{\text{first digit is 6 or greater}\}$$

From the table of probabilities,

$$P(A) = P(1) = 0.301$$
$$P(B) = P(6) + P(7) + P(8) + P(9)$$
$$= 0.067 + 0.058 + 0.051 + 0.046 = 0.222$$

Note that $P(B)$ is not the same as the probability that a random digit is greater than 6. The probability $P(6)$ that a first digit is 6 is included in "6 or greater" but not in "greater than 6."

The probability that a first digit is anything other than a 1 is, by the complement rule,

$$P(A^c) = 1 - P(A)$$
$$= 1 - 0.301 = 0.699$$

The events A and B are disjoint, so the probability that a first digit either is 1 or is 6 or greater is, by the addition rule,

$$P(A \text{ or } B) = P(A) + P(B)$$
$$= 0.301 + 0.222 = 0.523$$

Be careful to apply the addition rule only to disjoint events. Check that the probability of the event C that a first digit is odd is

$$P(C) = P(1) + P(3) + P(5) + P(7) + P(9) = 0.609$$

The probability

$$P(B \text{ or } C) = P(1) + P(3) + P(5) + P(6) + P(7) + P(8) + P(9) = 0.727$$

is *not* the sum of $P(B)$ and $P(C)$, because events B and C are not disjoint. Outcomes 7 and 9 are common to both events.

Assigning probabilities: equally likely outcomes

Assigning correct probabilities to individual outcomes often requires long observation of the random phenomenon. In some special circumstances, however, we are willing to assume that individual outcomes are equally likely because of some balance in the phenomenon. Ordinary coins have a physical balance that should make heads and tails equally likely, for example, and the table of random digits comes from a deliberate randomization.

EXAMPLE 6.11 RANDOM DIGITS

You might think that first digits are distributed "at random" among the digits 1 to 9. The 9 possible outcomes would then be equally likely. The sample space for a single digit is

$$S = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

Because the total probability must be 1, the probability of each of the 9 outcomes must be 1/9. That is, the assignment of probabilities to outcomes is

First digit:	1	2	3	4	5	6	7	8	9
Probability:	1/9	1/9	1/9	1/9	1/9	1/9	1/9	1/9	1/9

The probability of the event B that a randomly chosen first digit is 6 or greater is

$$\begin{aligned} P(B) &= P(6) + P(7) + P(8) + P(9) \\ &= \frac{1}{9} + \frac{1}{9} + \frac{1}{9} + \frac{1}{9} = \frac{4}{9} = 0.444 \end{aligned}$$

Compare this with the Benford's Law probability in Example 6.10. A crook who fakes data by using "random" digits will end up with too many first digits 6 or greater and too few 1s and 2s.

In Example 6.11 all outcomes have the same probability. Because there are 9 equally likely outcomes, each must have probability 1/9. Because exactly 4 of the 9 equally likely outcomes are 6 or greater, the probability of this event is 4/9. In the special situation where all outcomes are equally likely, we have a simple rule for assigning probabilities to events.

EQUALLY LIKELY OUTCOMES

If a random phenomenon has k possible outcomes, all equally likely, then each individual outcome has probability $1/k$. The probability of any event A is

$$\begin{aligned} P(A) &= \frac{\text{count of outcomes in A}}{\text{count of outcomes in S}} \\ &= \frac{\text{count of outcomes in A}}{k} \end{aligned}$$

Most random phenomena do not have equally likely outcomes, so the general rule for finite sample spaces is more important than the special rule for equally likely outcomes.

EXERCISES

6.19 BLOOD TYPES All human blood can be typed as one of O, A, B, or AB, but the distribution of the types varies a bit with race. Here is the distribution of the blood type of a randomly chosen black American:

Blood type:	O	A	B	AB
Probability:	0.49	0.27	0.20	?

(a) What is the probability of type AB blood? Why?

(b) Maria has type B blood. She can safely receive blood transfusions from people with blood types O and B. What is the probability that a randomly chosen black American can donate blood to Maria?

6.20 DISTRIBUTION OF M&M COLORS If you draw an M&M candy at random from a bag of the candies, the candy you draw will have one of six colors. The probability of drawing each color depends on the proportion of each color among all candies made.

(a) The table below gives the probability of each color for a randomly chosen plain M&M:

Color:	Brown	Red	Yellow	Green	Orange	Blue
Probability:	0.3	0.2	0.2	0.1	0.1	?

What must be the probability of drawing a blue candy?

(b) The probabilities for peanut M&Ms are a bit different. Here they are:

Color:	Brown	Red	Yellow	Green	Orange	Blue
Probability:	0.2	0.1	0.2	0.1	0.1	?

What is the probability that a peanut M&M chosen at random is blue?

(c) What is the probability that a plain M&M is any of red, yellow, or orange? What is the probability that a peanut M&M has one of these colors?

6.21 HEART DISEASE AND CANCER Government data assign a single cause for each death that occurs in the United States. The data show that the probability is 0.45 that a randomly chosen death was due to cardiovascular (mainly heart) disease, and 0.22 that it was due to cancer. What is the probability that a death was due either to cardiovascular disease or to cancer? What is the probability that the death was due to some other cause?

6.22 DO HUSBANDS DO THEIR SHARE? The *New York Times* (August 21, 1989) reported a poll that interviewed a random sample of 1025 women. The married women in the sample were asked whether their husbands did their fair share of household chores. Here are the results:

Outcome	Probability
Does more than his fair share	0.12
Does his fair share	0.61
Does less than his fair share	?

These proportions are probabilities for the random phenomenon of choosing a married woman at random and asking her opinion.

(a) What must be the probability that the woman chosen says that her husband does less than his fair share? Why?

(b) The event "I think my husband does at least his fair share" contains the first two outcomes. What is its probability?

6.23 **ACADEMIC RANK** Select a first-year college student at random and ask what his or her academic rank was in high school. Here are the probabilities, based on proportions from a large sample survey of first-year students:

Rank:	Top 20%	Second 20%	Third 20%	Fourth 20%	Lowest 20%
Probability:	0.41	0.23	0.29	0.06	0.01

(a) What is the sum of these probabilities? Why do you expect the sum to have this value?

(b) What is the probability that a randomly chosen first-year college student was not in the top 20% of his or her high school class?

(c) What is the probability that a first-year student was in the top 40% in high school?

6.24 **SPIN 123** Refer to the experiment described in Activity 6 and Exercise 6.16 (page 341).

(a) Determine the theoretical probability that at least one digit will occur in its correct place.

(b) Compare the theoretical probability with your experimental (empirical) results.

6.25 **TETRAHEDRAL DICE** Psychologists sometimes use tetrahedral dice to study our intuition about chance behavior. A tetrahedron is a pyramid (think of Egypt) with four identical faces, each a triangle with all sides equal in length. Label the four faces of a tetrahedral die with 1, 2, 3, and 4 spots.

(a) Give a probability model for rolling such a die and recording the number of spots on the down-face. Explain why you think your model is at least close to correct.

(b) Give a probability model for rolling two such dice. That is, write down all possible outcomes and give a probability to each. What is the probability that the sum of the down-faces is 5?

6.26 **BENFORD'S LAW** Example 6.10 (page 345) states that the first digits of numbers in legitimate records often follow a distribution known as Benford's Law. Here is the distribution:

First digit:	1	2	3	4	5	6	7	8	9
Probability:	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046

It was shown in Example 6.10 that

$$P(A) = P(\text{first digit is 1}) = 0.301$$
$$P(B) = P(\text{first digit is 6 or greater}) = 0.222$$
$$P(C) = P(\text{first digit is odd}) = 0.609$$

We will define event D to be {first digit is less than 4}. Using the union and intersection notation, find the following probabilities.

(a) $P(D)$
(b) $P(B \cup D)$
(c) $P(D^c)$
(d) $P(C \cap D)$
(e) $P(B \cap C)$

Independence and the multiplication rule

Rule 4, the addition rule for disjoint events, describes the probability that *one or the other* of two events A and B will occur in the special situation when A and B cannot occur together because they are disjoint. Now we will describe the probability that *both* events A and B occur, again only in a special situation. More general rules appear in Section 6.3.

Suppose that you toss a balanced coin twice. You are counting heads, so two events of interest are

$$A = \text{first toss is a head}$$
$$B = \text{second toss is a head}$$

The events A and B are not disjoint. They occur together whenever both tosses give heads. We want to compute the probability of the event {A and B} that *both* tosses are heads. The Venn diagram in Figure 6.6 illustrates the event {A and B} as the overlapping area that is common to both A and B.

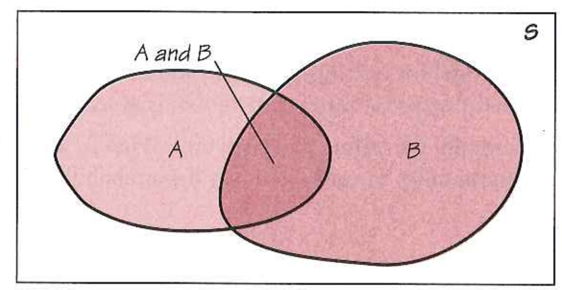


FIGURE 6.6 Venn diagram showing the event {*A* and *B*}.

The coin tossing of Buffon, Pearson, and Kerrich described at the beginning of this chapter makes us willing to assign probability 1/2 to a head when we toss a coin. So

$$P(A) = 0.5$$
$$P(B) = 0.5$$

What is $P(A \text{ and } B)$? Our common sense says that it is 1/4. The first toss will give a head half the time and then the second will give a head on half of those trials, so both tosses will give heads on $1/2 \times 1/2 = 1/4$ of all trials in the long run. This reasoning assumes that the second toss still has probability 1/2 of a head after the first has given a head. This is true—we can verify it by performing many trials of two tosses and observing the proportion of heads on the second toss after the first toss has produced a head. We say that the events "head on the first toss" and "head on the second toss" are **independent.** Here is our next probability rule.

independent

THE MULTIPLICATION RULE FOR INDEPENDENT EVENTS

Rule 5. Two events A and B are **independent** if knowing that one occurs does not change the probability that the other occurs. If A and B are independent,

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

This is the **multiplication rule** for independent events.

Our definition of independence is rather informal. A more precise definition appears in Section 6.3. In practice, though, we rarely need a precise definition of independence, because independence is usually *assumed* as part of a probability model when we want to describe random phenomena that seem to be physically unrelated to each other.

EXAMPLE 6.12 INDEPENDENT OR NOT INDEPENDENT?

Because a coin has no memory and most coin tossers cannot influence the fall of the coin, it is safe to assume that successive coin tosses are independent. For a balanced coin this means that after we see the outcome of the first toss, we still assign probability 1/2 to heads on the second toss.

On the other hand, the colors of successive cards dealt from the same deck are not independent. A standard 52-card deck contains 26 red and 26 black cards. For the first card dealt from a shuffled deck, the probability of a red card is $26/52 = 0.50$ because the 52 possible cards are equally likely. Once we see that the first card is red, we know that there are only 25 reds among the remaining 51 cards. The probability that the second card is red is therefore only $25/51 = 0.49$. Knowing the outcome of the first deal changes the probability for the second.

If a doctor measures your blood pressure twice, it is reasonable to assume that the two results are independent because the first result does not influence the instrument that makes the second reading. But if you take an IQ test or other mental test twice in succession, the two test scores are not independent. The learning that occurs on the first attempt influences your second attempt.

When independence is part of a probability model, the multiplication rule applies. Here is an example.

EXAMPLE 6.13 MENDEL'S PEAS

Gregor Mendel used garden peas in some of the experiments that revealed that inheritance operates randomly. The seed color of Mendel's peas can be either green or yellow. Two parent plants are "crossed" (one pollinates the other) to produce seeds. Each parent plant carries two genes for seed color, and each of these genes has probability 1/2 of being passed to a seed. The two genes that the seed receives, one from each parent, determine its color. The parents contribute their genes independently of each other.

Suppose that both parents carry the G and the Y genes. The seed will be green if both parents contribute a G gene; otherwise it will be yellow. If M is the event that the male contributes a G gene and F is the event that the female contributes a G gene, then the probability of a green seed is

$$\begin{aligned} P(M \text{ and } F) &= P(M)P(F) \\ &= (0.5)(0.5) = 0.25 \end{aligned}$$

In the long run, 1/4 of all seeds produced by crossing these plants will be green.

The multiplication rule $P(A \text{ and } B) = P(A)P(B)$ holds if A and B are independent but not otherwise. The addition rule $P(A \text{ or } B) = P(A) + P(B)$ holds if A and B are disjoint but not otherwise. Resist the temptation to use these simple formulas when the circumstances that justify them are not present. You must also be certain not to confuse disjointness and independence. If A and B are disjoint, then the fact that A occurs tells us that B cannot occur—look again at Figure 6.4. So **disjoint events are not independent.** Unlike disjointness or complements, independence cannot be pictured by a Venn diagram, because it involves the probabilities of the events rather than just the outcomes that make up the events.

Applying the probability rules

If two events A and B are independent, then their complements A^c and B^c are also independent and A^c is independent of B. Suppose, for example, that 75% of all registered voters in a suburban district are Republicans. If an opinion poll interviews two voters chosen independently, the probability that the first is a Republican and the second is not a Republican is (0.75)(0.25) = 0.1875. The multiplication rule also extends to collections of more than two events, provided that all are independent. Independence of events A, B, and C means that no information about any one or any two can change the probability of the remaining events. The formal definition is a bit messy. Fortunately, independence is usually assumed in setting up a probability model. We can then use the multiplication rule freely, as in this example.

EXAMPLE 6.14 ATLANTIC TELEPHONE CABLE

The first successful transatlantic telegraph cable was laid in 1866. The first telephone cable across the Atlantic did not appear until 1956—the barrier was designing "repeaters," amplifiers needed to boost the signal, that could operate for years on the sea bottom. This first cable had 52 repeaters. The copper cable, laid in 1983 and retired in 1994, had 662 repeaters. The first fiber optic cable was laid in 1988 and has 109 repeaters. There are now more than 400,000 miles of undersea cable, with more being laid every year to handle the flood of Internet traffic.

Repeaters in undersea cables must be very reliable. To see why, suppose that each repeater has probability 0.999 of functioning without failure for 25 years. Repeaters fail independently of each other. (This assumption means that there are no "common causes" such as earthquakes that would affect several repeaters at once.) Denote by A_i the event that the ith repeater operates successfully for 25 years.

The probability that two repeaters both last 25 years is

$$\begin{aligned} P(A_1 \text{ and } A_2) &= P(A_1)P(A_2) \\ &= 0.999 \times 0.999 = 0.998 \end{aligned}$$

For a cable with 10 repeaters the probability of no failures in 25 years is

$$\begin{aligned} P(A_1 \text{ and } A_2 \text{ and} \ldots \text{and } A_{10}) &= P(A_1)P(A_2) \cdots P(A_{10}) \\ &= 0.999 \times 0.999 \times \cdots \times 0.999 \\ &= 0.999^{10} = 0.990 \end{aligned}$$

Cables with 2 or 10 repeaters would be quite reliable. Unfortunately, the last copper transatlantic cable had 662 repeaters. The probability that all 662 work for 25 years is

$$P(A_1 \text{ and } A_2 \text{ and} \ldots \text{and } A_{662}) = 0.999^{662} = 0.516$$

This cable will fail to reach its 25-year design life about half the time if each repeater is 99.9% reliable over that period. The multiplication rule for probabilities shows that repeaters must be much more than 99.9% reliable.

By combining the rules we have learned, we can compute probabilities for rather complex events. Here is an example.

EXAMPLE 6.15 AIDS TESTING

Screening large numbers of blood samples for HIV, the virus that causes AIDS, uses an enzyme immunoassay (EIA) test that detects antibodies to the virus. Samples that test positive are retested using a more accurate "western blots" test. Applied to people who have no HIV antibodies, EIA has probability about 0.006 of producing a false positive. If the 140 employees of a medical clinic are tested and all 140 are free of HIV antibodies, what is the probability that at least one false positive will occur?

It is reasonable to assume as part of the probability model that the test results for different individuals are independent. The probability that the test is positive for a single person is 0.006, so the probability of a negative result is $1 - 0.006 = 0.994$ by the complement rule. The probability of at least one false positive among the 140 people tested is therefore

$$\begin{aligned} P(\text{at least one positive}) &= 1 - P(\text{no positives}) \\ &= 1 - P\,(140 \text{ negatives}) \\ &= 1 - 0.994^{140} \\ &= 1 - 0.431 = 0.569 \end{aligned}$$

The probability is greater than 1/2 that at least one of the 140 people will test positive for HIV even though no one has the virus.

EXERCISES

6.27 A BATTLE PLAN A general can plan a campaign to fight one major battle or three small battles. He believes that he has probability 0.6 of winning the large battle and probability 0.8 of winning each of the small battles. Victories or defeats in the small battles are independent. The general must win either the large battle or all three small battles to win the campaign. Which strategy should he choose?

6.28 DEFECTIVE CHIPS An automobile manufacturer buys computer chips from a supplier. The supplier sends a shipment containing 5% defective chips. Each chip chosen from this shipment has probability 0.05 of being defective, and each automobile uses 12 chips selected independently. What is the probability that all 12 chips in a car will work properly?

6.29 COLLEGE-EDUCATED LABORERS? Government data show that 26% of the civilian labor force have at least 4 years of college and that 16% of the labor force work as laborers or operators of machines or vehicles. Can you conclude that because $(0.26)(0.16) = 0.0416$, about 4% of the labor force are college-educated laborers or operators? Explain your answer.

6.30 Choose at random a U.S. resident at least 25 years of age. We are interested in the events

$$A = \{\text{The person chosen completed 4 years of college}\}$$
$$B = \{\text{The person chosen is 55 years old or older}\}$$

Government data recorded in Table 4.6 on page 241 allow us to assign probabilities to these events.

(a) Explain why $P(A) = 0.256$.

(b) Find $P(B)$.

(c) Find the probability that the person chosen is at least 55 years old *and* has 4 years of college education, $P(A \text{ and } B)$. Are the events A and B independent?

6.31 **BRIGHT LIGHTS?** A string of Christmas lights contains 20 lights. The lights are wired in series, so that if any light fails the whole string will go dark. Each light has probability 0.02 of failing during a 3-year period. The lights fail independently of each other. What is the probability that the string of lights will remain bright for 3 years?

6.32 **DETECTING STEROIDS** An athlete suspected of having used steroids is given two tests that operate independently of each other. Test A has probability 0.9 of being positive if steroids have been used. Test B has probability 0.8 of being positive if steroids have been used. What is the probability that *neither* test is positive if steroids have been used?

6.33 **TELEPHONE SUCCESS** Most sample surveys use random digit dialing equipment to call residential telephone numbers at random. The telephone polling firm Zogby International reports that the probability that a call reaches a live person is 0.2.[4] Calls are independent.

(a) A polling firm places 5 calls. What is the probability that none of them reaches a person?

(b) When calls are made to New York City, the probability of reaching a person is only 0.08. What is the probability that none of 5 calls made to New York City reaches a person?

SUMMARY

A **random phenomenon** has outcomes that we cannot predict but that nonetheless have a regular distribution in very many repetitions.

The **probability** of an event is the proportion of times the event occurs in many repeated trials of a random phenomenon.

A **probability model** for a random phenomenon consists of a sample space S and an assignment of probabilities P.

The **sample space S** is the set of all possible outcomes of the random phenomenon. Sets of outcomes are called **events.** P assigns a number $P(A)$ to an event A as its probability.

The **complement** A^c of an event A consists of exactly the outcomes that are not in A. Events A and B are **disjoint** (mutually exclusive) if they have no outcomes in common. Events A and B are **independent** if knowing that one event occurs does not change the probability we would assign to the other event.

Any assignment of probability must obey the rules that state the basic properties of probability:

1. $0 \le P(A) \le 1$ for any event A.
2. $P(S) = 1$.

3. **Complement rule:** For any event A, $P(A^c) = 1 - P(A)$.

4. **Addition rule:** If events A and B are **disjoint**, then $P(A \text{ or } B) = P(A \cup B) = P(A) + P(B)$.

5. **Multiplication rule:** If events A and B are **independent**, then $P(A \text{ and } B) = P(A \cap B) = P(A)P(B)$.

SECTION 6.2 EXERCISES

6.34 LEGITIMATE PROBABILITY MODEL? Figure 6.7 displays several assignments of probabilities to the six faces of a die. We can learn which assignment is actually *accurate* for a particular die only by rolling the die many times. However, some of the assignments are not *legitimate* assignments of probability. That is, they do not obey the rules. Which are legitimate and which are not? In the case of the illegitimate models, explain what is wrong.

Outcome	Model 1	Model 2	Model 3	Model 4
⚀	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{3}$
⚁	0	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{3}$
⚂	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{7}$	$-\frac{1}{6}$
⚃	0	$\frac{1}{6}$	$\frac{1}{7}$	$-\frac{1}{6}$
⚄	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{3}$
⚅	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{3}$

FIGURE 6.7 Four assignments of probabilities to the six faces of a die.

6.35 LEGITIMATE ASSIGNMENT OF PROBABILITIES? In each of the following situations, state whether or not the given assignment of probabilities to individual outcomes is legitimate, that is, satisfies the rules of probability. If not, give specific reasons for your answer.

(a) When a coin is spun, $P(H) = 0.55$ and $P(T) = 0.45$.

(b) When two coins are tossed, $P(HH) = 0.4$, $P(HT) = 0.4$, $P(TH) = 0.4$, and $P(TT) = 0.4$.

(c) When a die is rolled, the number of spots on the up-face has $P(1) = 1/2$, $P(4) = 1/6$, $P(5) = 1/6$, and $P(6) = 1/6$.

6.36 CAR COLORS Choose a new car or light truck at random and note its color. Here are the probabilities of the most popular colors for vehicles made in North America in 2000:[5]

Color:	Silver	White	Black	Dark green	Dark blue	Medium red
Probability:	0.176	0.172	0.113	0.089	0.088	0.067

(a) What is the probability that the vehicle you choose has any color other than the six listed?

(b) What is the probability that a randomly chosen vehicle is either silver or white?

(c) Choose two vehicles at random. What is the probability that both are silver or white?

6.37 NEW CENSUS CATEGORIES The 2000 census allowed each person to choose one or more from a long list of races. That is, in the eyes of the Census Bureau, you belong to whatever race or races you say you belong to. "Hispanic/Latino" is a separate category; Hispanics may be of any race. If we choose a resident of the United States at random, the 2000 census gives these probabilities:

	Hispanic	Not Hispanic
Asian	0.000	0.036
Black	0.003	0.121
White	0.060	0.691
Other	0.062	0.027

Let A be the event that a randomly chosen American is Hispanic, and let B be the event that the person chosen is white.

(a) Verify that the table gives a legitimate assignment of probabilities.

(b) What is $P(A)$?

(c) Describe B^c in words and find $P(B^c)$ by the complement rule.

(d) Express "the person chosen is a non-Hispanic white" in terms of events A and B. What is the probability of this event?

6.38 BEING HISPANIC Exercise 6.37 assigns probabilities for the ethnic background of a randomly chosen resident of the United States. Let A be the event that the person chosen is Hispanic, and let B be the event that he or she is white. Are events A and B independent? How do you know?

6.39 PREPARING FOR THE GMAT A company that offers courses to prepare would-be M.B.A. students for the GMAT examination finds that 40% of its customers are currently undergraduate students and 60% are college graduates. After completing the course, 50% of the undergraduates and 70% of the graduates achieve scores of at least 600 on the GMAT.

(a) What proportion of customers are undergraduates *and* score at least 600? What proportion of customers are graduates *and* score at least 600?

(b) What proportion of all customers score at least 600 on the GMAT?

6.40 THE RISE AND FALL OF PORTFOLIO VALUES The "random walk" theory of securities prices holds that price movements in disjoint time periods are independent of each other. Suppose that we record only whether the price is up or down each year, and that the probability that our portfolio rises in price in any one year is 0.65. (This probability is approximately correct for a portfolio containing equal dollar amounts of all common stocks listed on the New York Stock Exchange.)

(a) What is the probability that our portfolio goes up for 3 consecutive years?

(b) If you know that the portfolio has risen in price 2 years in a row, what probability do you assign to the event that it will go down next year?

(c) What is the probability that the portfolio's value moves in the same direction in both of the next 2 years?

6.41 USING A TABLE TO FIND PROBABILITIES The type of medical care a patient receives may vary with the age of the patient. A large study of women who had a breast lump investigated whether or not each woman received a mammogram and a biopsy when the lump was discovered. Here are some probabilities estimated by the study. The entries in the table are the probabilities that *both* of two events occur; for example, 0.321 is the probability that a patient is under 65 years of age *and* the tests were done. The four probabilities in the table have sum 1 because the table lists all possible outcomes.

	Tests done?	
	Yes	No
Age under 65:	0.321	0.124
Age 65 or over:	0.365	0.190

(a) What is the probability that a patient in this study is under 65? That a patient is 65 or over?

(b) What is the probability that the tests were done for a patient? That they were not done?

(c) Are the events A = {patient was 65 or older} and B = {the tests were done} independent? Were the tests omitted on older patients more or less frequently than would be the case if testing were independent of age?

6.42 ROULETTE A roulette wheel has 38 slots, numbered 0, 00, and 1 to 36. The slots 0 and 00 are colored green, 18 of the others are red, and 18 are black. The dealer spins the wheel and at the same time rolls a small ball along the wheel in the opposite direction. The wheel is carefully balanced so that the ball is equally likely to land in any slot when the wheel slows. Gamblers can bet on various combinations of numbers and colors.

(a) What is the probability that the ball will land in any one slot?

(b) If you bet on "red," you win if the ball lands in a red slot. What is the probability of winning?

(c) The slot numbers are laid out on a board on which gamblers place their bets. One column of numbers on the board contains all multiples of 3, that is, 3, 6, 9, . . . , 36. You place a "column bet" that wins if any of these numbers comes up. What is your probability of winning?

6.43 WHICH IS MOST LIKELY? A six-sided die has four green and two red faces and is balanced so that each face is equally likely to come up. The die will be rolled several times. You must choose one of the following three sequences of colors; you will win $25 if the first rolls of the die give the sequence that you have chosen.

RGRRR
RGRRRG
GRRRRR

Which sequence do you choose? Explain your choice.[6]

6.44 ALBINISM IN GENETICS The gene for albinism in humans is recessive. That is, carriers of this gene have probability 1/2 of passing it to a child, and the child is albino only if both parents pass the albinism gene. Parents pass their genes independently of each other. If both parents carry the albinism gene, what is the probability that their first child is albino? If they have two children (who inherit independently of each other), what is the probability that both are albino? That neither is albino?

6.45 DISJOINT VERSUS INDEPENDENT EVENTS This exercise explores the relationship between mutually exclusive and independent events.

(a) Assume that events A and B are non-empty, independent events. Show that A and B must intersect (i.e., that $A \cap B \neq \varnothing$).

(b) Use the results of (a) to argue that if A and B are disjoint, then they cannot be independent.

(c) Find an example of two events that are neither disjoint nor independent.

6.3 GENERAL PROBABILITY RULES

In this section we will consider some additional laws that govern any assignment of probabilities. The purpose of learning more laws of probability is to be able to give probability models for more complex random phenomena. We have already met and used five rules.

RULES OF PROBABILITY

Rule 1. $0 \le P(A) \le 1$ for any event A.

Rule 2. $P(S) = 1$.

Rule 3. **Complement rule:** For any event A,

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

Rule 4. **Addition rule:** If A and B are **disjoint** events, then

$$P(A \text{ or } B) = P(A) + P(B)$$

Rule 5. **Multiplication rule:** If A and B are **independent** events, then

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

General addition rules

Probability has the property that if A and B are disjoint events, then $P(A \text{ or } B) = P(A) + P(B)$. What if there are more than two events, or if the events are not disjoint? These circumstances are covered by more general addition rules for probability.

UNION

The **union** of any collection of events is the event that at least one of the collection occurs.

For two events A and B, the union is the event $\{A \text{ or } B\}$ that A or B or both occur. From the addition rule for two disjoint events, we can obtain rules for more general unions. Suppose first that we have several events—say A, B, and C—that are disjoint in pairs. That is, no two can occur simultaneously. The Venn diagram in Figure 6.8 illustrates three disjoint events.

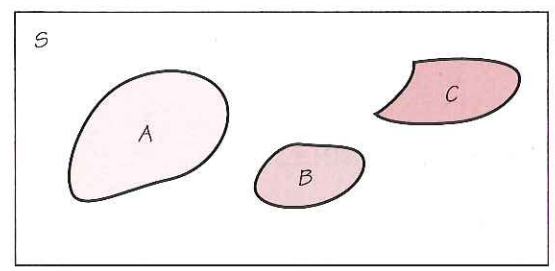


FIGURE 6.8 The addition rule for disjoint events: $P(A \text{ or } B \text{ or } C) = P(A) + P(B) + P(C)$ when events A, B, and C are disjoint.

The addition rule for two disjoint events extends to the following law.

ADDITION RULE FOR DISJOINT EVENTS

If events A, B, and C are disjoint in the sense that no two have any outcomes in common, then

$$P(\text{one or more of } A, B, C) = P(A) + P(B) + P(C)$$

This rule extends to any number of disjoint events.

EXAMPLE 6.16 UNIFORM DISTRIBUTION

Generate a random number X between 0 and 1. What is the probability that the first digit will be odd? We will learn in Chapter 7 that the variable X has the density curve of a uniform distribution (see Exercise 2.2, page 83.). This density curve has constant height 1 between 0 and 1 and is 0 elsewhere. The event that the first digit of X is odd is the union of five disjoint events. These events are

$$0.10 \le X < 0.20$$
$$0.30 \le X < 0.40$$
$$0.50 \le X < 0.60$$
$$0.70 \le X < 0.80$$
$$0.90 \le X < 1.00$$

Figure 6.9 illustrates the probabilities of these events as areas under the density curve. Each has probability 0.1 equal to its length. The union of the five therefore has probability equal to the sum, or 0.5. As we should expect, a random number is equally likely to begin with an odd or an even digit.

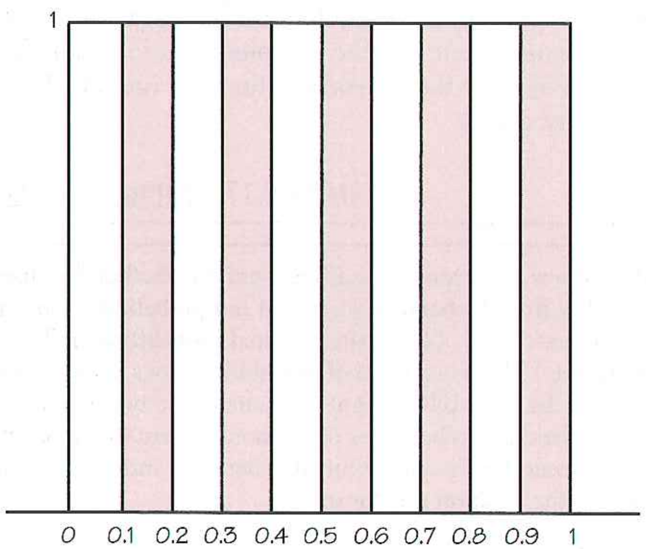


FIGURE 6.9 The probability that the first digit of a random number is odd is the sum of the probabilities of the 5 disjoint events shown.

If events A and B are *not* disjoint, they can occur simultaneously. The probability of their union is then *less* than the sum of their probabilities. As Figure 6.10 suggests, the outcomes common to both are counted twice when we add probabilities, so we must subtract this probability once.

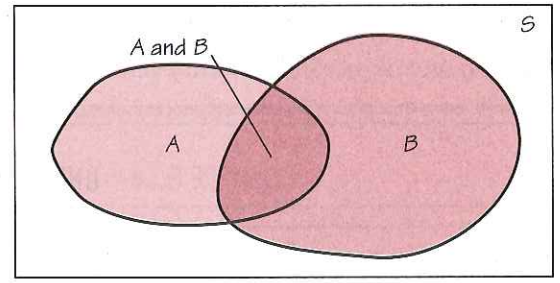


FIGURE 6.10 The general addition rule for the union of two events: $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$ for any events A and B.

Here is the addition rule for the union of any two events, disjoint or not.

GENERAL ADDITION RULE FOR UNIONS OF TWO EVENTS

For any two events A and B,

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

Equivalently,

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

If A and B are disjoint, the event {A and B} that both occur has no outcomes in it. This *empty event* $\varnothing$ is the complement of the sample space S and must have probability 0. So the general addition rule includes Rule 4, the addition rule for disjoint events.

EXAMPLE 6.17 PROBABILITY OF PROMOTION

Deborah and Matthew are anxiously awaiting word on whether they have been made partners of their law firm. Deborah guesses that her probability of making partner is 0.7 and that Matthew's is 0.5. (These are personal probabilities reflecting Deborah's assessment of chance.) This assignment of probabilities does not give us enough information to compute the probability that at least one of the two is promoted. In particular, adding the individual probabilities of promotion gives the impossible result 1.2. If Deborah also guesses that the probability that *both* she and Matthew are made partners is 0.3, then by the addition rule for unions

$$P(\text{at least one is promoted}) = 0.7 + 0.5 - 0.3 = 0.9$$

The probability that *neither* is promoted is then 0.1 by the complement rule.

Venn diagrams are a great help in finding probabilities for unions, because you can just think of adding and subtracting areas. Figure 6.11 shows some events and their probabilities for Example 6.17. What is the probability that Deborah is promoted and Matthew is not? The Venn diagram shows that this is the probability that Deborah is promoted minus the probability that both are promoted, 0.7 – 0.3 = 0.4. Similarly, the probability that Matthew is promoted and Deborah is not is 0.5 – 0.3 = 0.2. The four probabilities that appear in the figure add to 1 because they refer to four disjoint events whose union is the entire sample space.

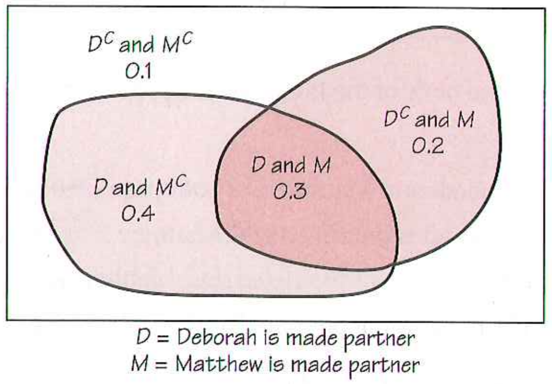


FIGURE 6.11 Venn diagram and probabilities.

The simultaneous occurrence of two events, such as A = Deborah is promoted *and* B = Matthew is promoted, is called a ***joint event.*** The probability of a joint event, such as P(Deborah is promoted *and* Matthew is promoted) = $P(A \text{ and } B)$, is called a ***joint probability.*** Determining joint probabilities when you have equally likely outcomes can be as easy as counting outcomes. For most situations, however, we will need more powerful methods, which will be developed later in this section.

Here's another way to work with joint events. We have two variables. One variable is employee, which has two values: Deborah and Matthew. The other variable is promotion, which also has two values: promoted and not promoted.

$$D = \{\text{Deborah promoted}\}$$
$$D^c = \{\text{Deborah not promoted}\}$$
$$M = \{\text{Matthew promoted}\}$$
$$M^c = \{\text{Matthew not promoted}\}$$

We can construct a table and write in the probabilities that Deborah assumes:

		Matthew		
		Promoted	Not promoted	Total
Deborah	Promoted	0.3		0.7
	Not promoted			
	Total	0.5		1

The rows and columns have to add to the totals shown, so we can fill in the rest of the table to produce the completed table:

		Matthew		
		Promoted	Not promoted	Total
Deborah	Promoted	0.3	0.4	0.7
	Not promoted	0.2	0.1	0.3
Total		0.5	0.5	1

The four entries in the body of the table are the probabilities of the joint events of interest:

$P(D \text{ and } M) = P(\text{Deborah and Matthew are both promoted}) = 0.3$

$P(D \text{ and } M^c) = P(\text{Deborah is promoted and Matthew is not promoted}) = 0.4$

$P(D^c \text{ and } M) = P(\text{Deborah is not promoted and Matthew is promoted}) = 0.2$

$P(D^c \text{ and } M^c) = P(\text{Deborah is not promoted and Matthew is not promoted}) = 0.1$

Note that these joint probabilities add to 1.

We will continue our discussion of tables and joint events in the next section.

EXERCISES

6.46 PROSPERITY AND EDUCATION Call a household prosperous if its income exceeds \$100,000. Call the household educated if the householder completed college. Select an American household at random, and let A be the event that the selected household is prosperous and B the event that it is educated. According to the Census Bureau, $P(A) = 0.134$, $P(B) = 0.254$, and the joint probability that a household is both prosperous and educated is $P(A \text{ and } B) = 0.080$. What is the probability $P(A \text{ or } B)$ that the household selected is either prosperous or educated?

6.47 Draw a Venn diagram that shows the relation between events A and B in Exercise 6.46. Indicate each of the following events on your diagram and use the information in Exercise 6.46 to calculate the probability of each event. Finally, describe in words what each event is.

(a) $\{A \text{ and } B\}$

(b) $\{A \text{ and } B^c\}$

(c) $\{A^c \text{ and } B\}$

(d) $\{A^c \text{ and } B^c\}$

6.48 WINNING CONTRACTS Consolidated Builders has bid on two large construction projects. The company president believes that the probability of winning the first contract (event A) is 0.6, that the probability of winning the second (event B) is 0.5, and that the joint probability of winning both jobs (event $\{A \text{ and } B\}$) is 0.3. What is the probability of the event $\{A \text{ or } B\}$ that Consolidated will win at least one of the jobs?

6.49 In the setting of the previous exercise, are events A and B independent? Do a calculation that proves your answer.

6.50 Draw a Venn diagram that illustrates the relation between events A and B in Exercise 6.48. Write each of the following events in terms of A, B, A^c, and B^c. Indicate the events on your diagram and use the information in Exercise 6.48 to calculate the probability of each.

(a) Consolidated wins both jobs.

(b) Consolidated wins the first job but not the second.

(c) Consolidated does not win the first job but does win the second.

(d) Consolidated does not win either job.

6.51 CAFFEINE IN THE DIET Common sources of caffeine are coffee, tea, and cola drinks. Suppose that

55% of adults drink coffee
25% of adults drink tea
45% of adults drink cola

and also that

15% drink both coffee and tea
5% drink all three beverages
25% drink both coffee and cola
5% drink only tea

Draw a Venn diagram marked with this information. Use it along with the addition rules to answer the following questions.

(a) What percent of adults drink only cola?

(b) What percent drink none of these beverages?

6.52 TASTES IN MUSIC Musical styles other than rock and pop are becoming more popular. A survey of college students finds that 40% like country music, 30% like gospel music, and 10% like both.

(a) Make a Venn diagram with these results.

(b) What percent of college students like country but not gospel?

(c) What percent like neither country nor gospel?

6.53 GETTING INTO COLLEGE Ramon has applied to both Princeton and Stanford. He thinks the probability that Princeton will admit him is 0.4, the probability that Stanford will admit him is 0.5, and the probability that both will admit him is 0.2.

(a) Make a Venn diagram with the probabilities given marked.

(b) What is the probability that neither university admits Ramon?

(c) What is the probability that he gets into Stanford but not Princeton?

Conditional probability

The probability we assign to an event can change if we know that some other event has occurred. This idea is the key to many applications of probability.

EXAMPLE 6.18 AMARILLO SLIM WANTS AN ACE

Slim is a professional poker player. He stares at the dealer, who prepares to deal. What is the probability that the card dealt to Slim is an ace? There are 52 cards in the deck. Because the deck was carefully shuffled, the next card dealt is equally likely to be any of the cards. Four of the 52 cards are aces. So

$$P(\text{ace}) = \frac{4}{52} = \frac{1}{13}$$

This calculation assumes that Slim knows nothing about any cards already dealt. Suppose now that he is looking at 4 cards already in his hand, and that 1 of them is an ace. He knows nothing about the other 48 cards except that exactly 3 aces are among them. Slim's probability of being dealt an ace, *given what he knows*, is now

$$P(\text{ace} \mid 1 \text{ ace in 4 visible cards}) = \frac{3}{48} = \frac{1}{16}$$

Knowing that there is one ace among the four cards Slim can see changes the probability that the next card dealt is an ace.

conditional probability

The new notation $P(A \mid B)$ is a ***conditional probability***. That is, it gives the probability of one event (the next card dealt is an ace) under the condition that we know another event (exactly one of the four visible cards is an ace). You can read the bar | as "given the information that."

In Example 6.18 we could find probabilities because we were willing to use an equally likely probability model for a shuffled deck of cards. Here is an example based on data.

EXAMPLE 6.19 MARITAL STATUS OF WOMEN

Table 6.1 shows the marital status of adult women broken down by age group.

TABLE 6.1 Age and marital status of women (thousands of women)

	Age			
	18–29	30–64	65 and over	Total
Married	7,842	43,808	8,270	59,920
Never married	13,930	7,184	751	21,865
Widowed	36	2,523	8,385	10,944
Divorced	704	9,174	1,263	11,141
Total	22,512	62,689	18,669	103,870

Source: Data for 1999 from the 2000 *Statistical Abstract of the United States.*

We are interested in the probability that a randomly chosen woman is married. It is common sense that knowing her age group will change the probability: many young women have not married, most middle-aged women are married, and older women are more likely to be widows. To help us think carefully, let's define two events:

$$A = \text{the woman chosen is young, ages 18 to 29}$$
$$B = \text{the woman chosen is married}$$

There are (in thousands) 103,870 adult women in the United States. Of these women, 22,512 are aged 18 to 29. Choosing at random gives each woman an equal chance, so the probability of choosing a young woman is

$$P(A) = \frac{22{,}512}{103{,}870} = 0.217$$

The table shows that there are 7842 thousand young married women. So the probability that we choose a woman who is both young and married is

$$P(A \text{ and } B) = \frac{7842}{103{,}870} = 0.075$$

To find the *conditional* probability that a woman is married *given the information* that she is young, look only at the "18–29" column. The young women are all in this column, so the information given says that only this column is relevant. The conditional probability is

$$P(B \mid A) = \frac{7842}{22{,}512} = 0.348$$

As we expected, the conditional probability that a woman is married when we know she is under age 30 is much higher than the probability for a randomly chosen woman.

It is easy to confuse the three probabilities in Example 6.19. Look carefully at Table 6.1 and be sure you understand the example. There is a relationship among these three probabilities. The probability that a woman is both young *and* married is the product of the probabilities that she is young and that she is married *given* that she is young. That is,

$$\begin{aligned} P(A \text{ and } B) &= P(A) \times P(B \mid A) \\ &= \frac{22{,}512}{103{,}870} \times \frac{7842}{22{,}512} \\ &= \frac{7842}{103{,}870} = 0.075 \qquad \text{(as before)} \end{aligned}$$

Try to think your way through this in words: First, the woman is young; then, given that she is young, she is married. We have just discovered the fundamental multiplication rule of probability.

GENERAL MULTIPLICATION RULE FOR ANY TWO EVENTS

The probability that both of two events A and B happen together can be found by

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

Here $P(B \mid A)$ is the conditional probability that B occurs given the information that A occurs.

In words, this rule says that for both of two events to occur, first one must occur and then, given that the first event has occurred, the second must occur. In our example, the joint probability that a randomly chosen woman is both age 18 to 29 (event A) and married (event B) is

$$\begin{aligned} P(A \text{ and } B) &= P(A)P(B \mid A) \\ &= (0.217)(0.348) = 0.076 \end{aligned}$$

EXAMPLE 6.20 SLIM WANTS DIAMONDS

Slim is still at the poker table. At the moment, he wants very much to draw 2 diamonds in a row. As he looks at his hand and at the upturned cards on the table, Slim sees 11 cards. Of these, 4 are diamonds. The full deck contains 13 diamonds among its 52 cards, so 9 of the 41 unseen cards are diamonds. To find Slim's probability of drawing two diamonds, first calculate

$$P(\text{first card diamond}) = \frac{9}{41}$$

$$P(\text{second card diamond} \mid \text{first card diamond}) = \frac{8}{40}$$

Slim finds both probabilities by counting cards. The probability that the first card drawn is a diamond is 9/41 because 9 of the 41 unseen cards are diamonds. If the first card is a diamond, that leaves 8 diamonds among the 40 remaining cards. So the *conditional* probability of another diamond is 8/40. The multiplication rule now says that

$$P(\text{both cards diamonds}) = \frac{9}{41} \times \frac{8}{40} = 0.044$$

Slim will need luck to draw his diamonds.

If we know $P(A)$ and $P(A \text{ and } B)$, we can rearrange the general multiplication rule to produce a *definition* of the conditional probability $P(B \mid A)$ in terms of unconditional probabilities.

DEFINITION OF CONDITIONAL PROBABILITY

When $P(A) > 0$, the conditional probability of B given A is

$$P(B \mid A) = \frac{P(A \text{ and } B)}{P(A)}$$

Be sure to keep in mind the distinct roles in $P(B \mid A)$ of the event B whose probability we are computing and the event A that represents the information we are given. The conditional probability $P(B \mid A)$ makes no sense if the event A can never occur, so we require that $P(A) > 0$ whenever we talk about $P(B \mid A)$.

EXAMPLE 6.21 FINDING CONDITIONAL PROBABILITIES

What is the conditional probability that a woman is a widow, given that she is at least 65 years old? We see from Table 6.1 that

$$P(\text{at least } 65) = \frac{18{,}669}{103{,}870} = 0.180$$

$$P(\text{widowed } \textit{and} \text{ at least } 65) = \frac{8385}{103{,}870} = 0.081$$

The conditional probability is therefore

$$P(\text{widowed} \mid \text{at least } 65) = \frac{P(\text{widowed } \textit{and} \text{ at least } 65)}{P(\text{at least } 65)}$$

$$= \frac{0.081}{0.180} = 0.450$$

Check that this agrees (up to roundoff error) with the result obtained from the "65 and over" column of Table 6.1:

$$P(\text{widowed} \mid \text{at least } 65) = \frac{8385}{18{,}669} = 0.449$$

EXERCISES

6.54 **AMERICAN WOMEN, I** Choose an adult American woman at random. Table 6.1 describes the population from which we draw. Use the information in that table to answer the following questions.

(a) What is the probability that the woman chosen is 65 years old or older?

(b) What is the conditional probability that the woman chosen is married, given that she is 65 or over?

(c) How many women are *both* married and in the over-65 age group? What is the probability that the woman we choose is a married woman at least 65 years old?

(d) Verify that the three probabilities you found in (a), (b), and (c) satisfy the multiplication rule.

6.55 AMERICAN WOMEN, II Choose an adult American woman at random. Table 6.1 describes the population from which we draw.

(a) What is the conditional probability that the woman chosen is 18 to 29 years old, given that she is married?

(b) In Example 6.19 we found that $P(\text{married} \mid \text{age 18 to 29}) = 0.348$. Complete this sentence: 0.348 is the proportion of women who are _____ among those women who are _____.

(c) In (a), you found $P(\text{age 18 to 29} \mid \text{married})$. Write a sentence of the form given in (b) that describes the meaning of this result. The two conditional probabilities give us very different information.

6.56 WOMAN MANAGERS Choose an employed person at random. Let A be the event that the person chosen is a woman, and B the event that the person holds a managerial or professional job. Government data tell us that $P(A) = 0.46$ and the probability of managerial and professional jobs among women is $P(B \mid A) = 0.32$. Find the probability that a randomly chosen employed person is a woman holding a managerial or professional position.

6.57 BUYING FROM JAPAN Functional Robotics Corporation buys electrical controllers from a Japanese supplier. The company's treasurer thinks that there is probability 0.4 that the dollar will fall in value against the Japanese yen in the next month. The treasurer also believes that *if* the dollar falls there is probability 0.8 that the supplier will demand renegotiation of the contract. What probability has the treasurer assigned to the event that the dollar falls and the supplier demands renegotiation?

6.58 THE PROBABILITY OF A FLUSH A poker player holds a flush when all 5 cards in the hand belong to the same suit. We will find the probability of a flush when 5 cards are dealt. Remember that a deck contains 52 cards, 13 of each suit, and that when the deck is well shuffled, each card dealt is equally likely to be any of those that remain in the deck.

(a) We will concentrate on spades. What is the probability that the first card dealt is a spade? What is the conditional probability that the second card is a spade, given that the first is a spade?

(b) Continue to count the remaining cards to find the conditional probabilities of a spade on the third, the fourth, and the fifth card, given in each case that all previous cards are spades.

(c) The probability of being dealt 5 spades is the product of the five probabilities you have found. Why? What is this probability?

(d) The probability of being dealt 5 hearts or 5 diamonds or 5 clubs is the same as the probability of being dealt 5 spades. What is the probability of being dealt a flush?

6.59 THE PROBABILITY OF A ROYAL FLUSH A royal flush is the highest hand possible in poker. It consists of the ace, king, queen, jack, and ten of the same suit. Modify the outline given in Exercise 6.58 to find the probability of being dealt a royal flush in a five-card deal.

6.60 INCOME TAX RETURNS Here is the distribution of the adjusted gross income (in thousands of dollars) reported on individual federal income tax returns in 1994:

Income:	<10	10–29	30–49	50–99	≥100
Probability:	0.12	0.39	0.24	0.20	0.05

(a) What is the probability that a randomly chosen return shows an adjusted gross income of $50,000 or more?

(b) Given that a return shows an income of at least $50,000, what is the conditional probability that the income is at least $100,000?

6.61 TASTES IN MUSIC Musical styles other than rock and pop are becoming more popular. A survey of college students finds that 40% like country music, 30% like gospel music, and 10% like both.

(a) What is the conditional probability that a student likes gospel music if we know that he or she likes country music?

(b) What is the conditional probability that a student who does not like country music likes gospel music? (A Venn diagram may help you.)

Extended multiplication rules

The definition of conditional probability reminds us that in principle all probabilities, including conditional probabilities, can be found from the assignment of probabilities to events that describe a random phenomenon. More often, however, conditional probabilities are part of the information given to us in a probability model, and the multiplication rule is used to compute $P(A \text{ and } B)$.

The union of a collection of events is the event that *any* of them occur. Here is the corresponding term for the event that *all* of them occur.

INTERSECTION

The **intersection** of any collection of events is the event that *all* of the events occur.

To extend the multiplication rule to the probability that all of several events occur, the key is to condition each event on the occurrence of *all* of the preceding events. For example, the intersection of three events A, B, and C has probability

$$P(A \text{ and } B \text{ and } C) = P(A)P(B \mid A)P(C \mid A \text{ and } B)$$

EXAMPLE 6.22 THE FUTURE OF HIGH SCHOOL ATHLETES

Only 5% of male high school basketball, baseball, and football players go on to play at the college level. Of these, only 1.7% enter major league professional sports. About 40% of the athletes who compete in college and then reach the pros have a career of more than 3 years.[7] Define these events:

$$A = \{\text{competes in college}\}$$
$$B = \{\text{competes professionally}\}$$
$$C = \{\text{pro career longer than 3 years}\}$$

What is the probability that a high school athlete competes in college and then goes on to have a pro career of more than 3 years? We know that

$$P(A) = 0.05$$
$$P(B \mid A) = 0.017$$
$$P(C \mid A \text{ and } B) = 0.4$$

The probability we want is therefore

$$\begin{aligned} P(A \text{ and } B \text{ and } C) &= P(A)P(B \mid A)P(C \mid A \text{ and } B) \\ &= 0.05 \times 0.017 \times 0.40 = 0.00034 \end{aligned}$$

Only about 3 of every 10,000 high school athletes can expect to compete in college and have a professional career of more than 3 years. High school students would be wise to concentrate on studies rather than on unrealistic hopes of fortune from pro sports.

Tree diagrams revisited

Probability problems often require us to combine several of the basic rules into a more elaborate calculation. Here is an example that illustrates how to solve problems that have several stages.

EXAMPLE 6.23 A FUTURE IN PROFESSIONAL SPORTS?

What is the probability that a male high school athlete will go on to professional sports? In the notation of Example 6.22, this is $P(B)$. To find $P(B)$ from the information in Example 6.22, use the tree diagram in Figure 6.12 to organize your thinking.

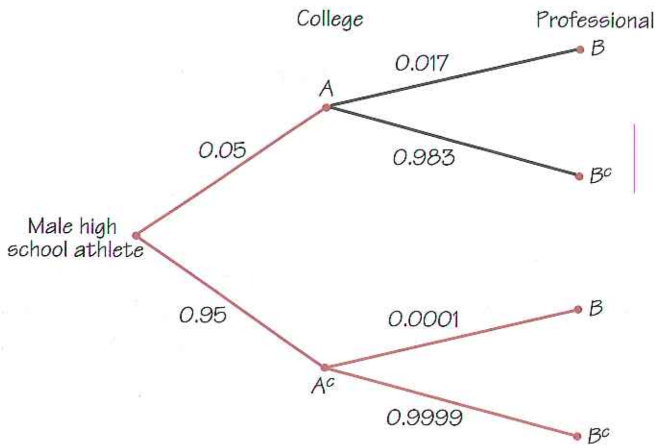


FIGURE 6.12 Tree diagram. The probability $P(B)$ is the sum of the probabilities of the two branches ending at B.

Each segment in the tree is one stage of the problem. Each complete branch shows a path that an athlete can take. The probability written on each segment is the conditional probability that an athlete follows that segment given that he has reached the point from which it branches. Starting at the left, high school athletes either do or do not compete in college. We know that the probability of competing in college is $P(A) = 0.05$, so the probability of not competing is $P(A^c) = 0.95$. These probabilities mark the leftmost branches in the tree.

Conditional on competing in college, the probability of playing professionally is $P(B \mid A) = 0.017$. So the conditional probability of *not* playing professionally is

$$P(B^c \mid A) = 1 - P(B \mid A) = 1 - 0.017 = 0.983$$

These conditional probabilities mark the paths branching out from A in Figure 6.12.

The lower half of the tree diagram describes athletes who do not compete in college (A^c). It is unusual for these athletes to play professionally, but a few go straight from high school to professional leagues. Suppose that the conditional probability that a high school athlete reaches professional play given that he does not compete in college is $P(B \mid A^c) = 0.0001$. We can now mark the two paths branching from A^c in Figure 6.12.

There are two disjoint paths to B (professional play). By the addition rule, $P(B)$ is the sum of their probabilities. The probability of reaching B through college (top half of the tree) is

$$\begin{aligned} P(B \text{ and } A) &= P(A)P(B \mid A) \\ &= 0.05 \times 0.017 = 0.00085 \end{aligned}$$

The probability of reaching B without college is

$$\begin{aligned} P(B \text{ and } A^c) &= P(A^c)P(B \mid A^c) \\ &= 0.95 \times 0.0001 = 0.000095 \end{aligned}$$

The final result is

$$P(B) = 0.00085 + 0.000095 = 0.000945$$

About 9 high school athletes out of 10,000 will play professional sports.

Tree diagrams combine the addition and multiplication rules. The multiplication rule says that **the probability of reaching the end of any complete branch is the product of the probabilities written on its segments.** The probability of any outcome, such as the event B that an athlete reaches professional sports, is then found by adding the probabilities of all branches that are part of that event.

Bayes's rule

There is another kind of probability question that we might ask in the context of studies of athletes. Our earlier calculations look forward toward professional sports as the final stage of an athlete's career. Now let's concentrate on professional athletes and look back at their earlier careers.

EXAMPLE 6.24 LOOKING BACK

What proportion of professional athletes competed in college? In the notation of Examples 6.22 and 6.23 this is the conditional probability $P(A \mid B)$. We start from the definition of conditional probability and then apply the results of Example 6.23:

$$\begin{aligned} P(A \mid B) &= \frac{P(A \text{ and } B)}{P(B)} \\ &= \frac{0.00085}{0.000945} = 0.8995 \end{aligned}$$

Almost 90% of professional athletes competed in college.

We know the probabilities $P(A)$ and $P(A^c)$ that a high school athlete does and does not compete in college. We also know the conditional probabilities $P(B \mid A)$ and $P(B \mid A^c)$ that an athlete from each group reaches professional sports. Example 6.23 shows how to use this information to calculate $P(B)$. The method can be summarized in a single expression that adds the probabilities of the two paths to B in the tree diagram:

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

In Example 6.24 we calculated the "reverse" conditional probability $P(A \mid B)$. The denominator 0.000945 in that example came from the expression just above. Put in this general notation, we have another probability law.

BAYES'S RULE

If A and B are any events whose probabilities are not 0 or 1,

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid A^c)P(A^c)}$$

Bayes's rule is named after Thomas Bayes, who wrestled with arguing from outcomes like B back to antecedents like A in a book published in 1763. It is far better to think your way through problems like Examples 6.23 and 6.24 rather than memorize these formal expressions.

Independence again

The conditional probability $P(B \mid A)$ is generally not equal to the unconditional probability $P(B)$. That is because the occurrence of event A generally gives us some additional information about whether or not event B occurs. If knowing that A occurs gives no additional information about B, then A and B are independent events. The formal definition of independence is expressed in terms of conditional probability.

INDEPENDENT EVENTS

Two events A and B that both have positive probability are **independent** if

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

This definition makes precise the informal description of independence given in Section 6.2. We now see that the multiplication rule for independent events, $P(A \text{ and } B) = P(A)P(B)$, is a special case of the general multiplication rule, $P(A \text{ and } B) = P(A)P(B \mid A)$, just as the addition rule for disjoint events is a special case of the general addition rule.

Decision analysis

One kind of decision making in the presence of uncertainty seeks to make the probability of a favorable outcome as large as possible. Here is an example that illustrates how the multiplication and addition rules, organized with the help of a tree diagram, apply to a decision problem.

EXAMPLE 6.25 TRANSPLANT OR DIALYSIS?

Lynn has end-stage kidney disease: her kidneys have failed so that she cannot survive unaided. Only about 52% of patients survive for 3 years with kidney dialysis. Fortunately, a kidney is available for transplant. Lynn's doctor gives her the following information for patients in her condition.

Transplant operations usually succeed. After 1 month, 96% of the transplanted kidneys are functioning. Three percent fail to function, and the patient must return to dialysis. The remaining 1% of the patients die within a month. Patients who return to dialysis have the same chance (52%) of surviving 3 years as if they had not attempted a transplant.

Of the successful transplants, however, only 82% continue to function for 3 years. Another 8% of these patients must return to dialysis, and 70% of these survive to the 3-year mark. The remaining 10% of "successful" patients die without returning to dialysis.[8]

There is too much information here to sort through without a tree diagram. The key is to realize that most of the percentages that Lynn's doctor gives her are conditional probabilities given that a patient has some specific prior history. Figure 6.13 is a tree diagram that organizes the information.

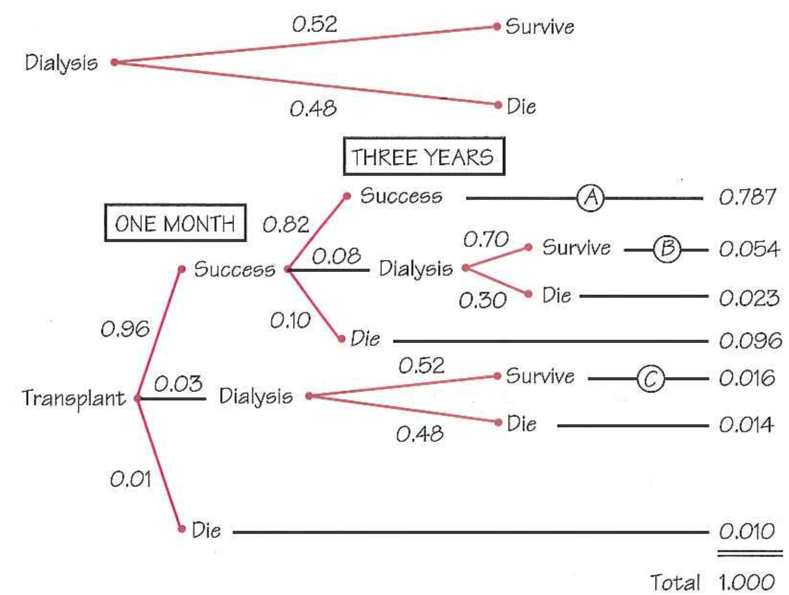


FIGURE 6.13 Tree diagram for the kidney failure decision problem.

Each path through the tree represents a possible outcome of Lynn's case. The probability written beside each branch after the first stage is the conditional probability of the next step given that Lynn has reached this point. For example, 0.82 is the conditional probability that a patient whose transplant succeeded survives 3 years with the transplant still functioning. The conditional probabilities of the other 3-year outcomes for a successful transplant are 0.08 and 0.10. They appear on the other branches from the "Success" node.

These three conditional probabilities add to 1 because these are all the possible outcomes following a successful transplant. Study the tree to convince yourself that it organizes all the information available.

The multiplication rule says that the probability of reaching the end of any path is the product of all the probabilities along that path. For example, look at the path marked A. The probability that a transplant succeeds and endures for 3 years is

$$P(\text{succeeds and lasts 3 years}) = P(\text{succeeds})P(\text{lasts 3 years} \mid \text{succeeds})$$
$$= (0.96)(0.82) = 0.787$$

Similarly, the path marked B is the event that a patient's transplant succeeds at the 1-month stage, fails before 3 years, and the patient nonetheless survives to 3 years after returning to dialysis. The probability of this is

$$P(B) = (0.96)(0.08)(0.70) = 0.054$$

The probabilities at the end of all the paths in Figure 6.13 add to 1 because these are all the possible 3-year outcomes.

What is the probability that Lynn will survive for 3 years if she has a transplant? This is the union of the three disjoint events marked A, B, and C in Figure 6.13. By the addition rule,

$$P(\text{survive}) = P(A) + P(B) + P(C)$$
$$= 0.787 + 0.054 + 0.016 = 0.857$$

Lynn's decision is easy: 0.857 is much higher than the probability 0.52 of surviving 3 years on dialysis. She will elect the transplant.

Where do the conditional probabilities in Example 6.25 come from? They are based in part on data—that is, on studies of many patients with kidney disease. But an individual's chances of survival depend on her age, general health, and other factors. Lynn's doctor considered her individual situation before giving her these particular probabilities. It is characteristic of most decision analysis problems that *personal probabilities* are used to describe the uncertainty of an informed decision maker.

EXERCISES

6.62 IRS RETURNS In 1999, the Internal Revenue Service received 127,075,145 individual tax returns. Of these, 9,534,653 reported an adjusted gross income of at least \$100,000 and 205,124 reported at least \$1 million.

(a) What is the probability that a randomly chosen individual tax return reports an income of at least \$100,000? At least \$1 million?

(b) If you know that the return chosen shows an income of \$100,000 or more, what is the conditional probability that the income is at least \$1 million?

6.63 SURGERY RISKS You have torn a tendon and are facing surgery to repair it. The orthopedic surgeon explains the risks to you. Infection occurs in 3% of such operations, the repair fails in 14%, and both infection and failure occur together in 1%. What percent of these operations succeed and are free from infection?

6.64 HIV TESTING Enzyme immunoassay (EIA) tests are used to screen blood specimens for the presence of antibodies to HIV, the virus that causes AIDS. Antibodies indicate the presence of the virus. The test is quite accurate but is not always correct. Here are approximate probabilities of positive and negative EIA outcomes when the blood tested does and does not actually contain antibodies to HIV.[9]

	Test result	
	+	−
Antibodies present:	0.9985	0.0015
Antibodies absent:	0.006	0.994

Suppose that 1% of a large population carries antibodies to HIV in their blood.

(a) Draw a tree diagram for selecting a person from this population (outcomes: antibodies present or absent) and for testing his or her blood (outcomes: EIA positive or negative).

(b) What is the probability that the EIA is positive for a randomly chosen person from this population?

(c) What is the probability that a person has the antibody given that the EIA test is positive?

(This exercise illustrates a fact that is important when considering proposals for widespread testing for HIV, illegal drugs, or agents of biological warfare: if the condition being tested is uncommon in the population, many positives will be false positives.)

6.65 The previous exercise gives data on the results of EIA tests for the presence of antibodies to HIV. Repeat part (c) of that exercise for two different populations:

(a) Blood donors are prescreened for HIV risk factors, so perhaps only 0.1% (0.001) of this population carries HIV antibodies.

(b) Clients of a drug rehab clinic are a high-risk group, so perhaps 10% of this population carries HIV antibodies.

(c) What general lesson do your calculations illustrate?

SUMMARY

The **complement** A^c of an event A contains all outcomes that are not in A. The **union** {A **or** B} of events A and B contains all outcomes in A, in B, or in both A and B. The **intersection** {A and B} contains all outcomes that are in both A and B, but not outcomes in A alone or B alone.

The essential general rules of elementary probability are

Legitimate values: $0 \le P(A) \le 1$ for any event A

Total probability 1: $P(S) = 1$

Complement rule: $P(A^c) = 1 - P(A)$

Addition rule: $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

Multiplication rule: $P(A \text{ and } B) = P(A)P(B \mid A)$

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

$$P(B \mid A) = \frac{P(A \text{ and } B)}{P(A)}$$

when $P(A) > 0$ but in practice is most often found from directly available information.

If A and B are **disjoint** (mutually exclusive), then $P(A \text{ and } B) = 0$. The general addition rule for unions then becomes the special addition rule, $P(A \text{ or } B) = P(A) + P(B)$.

A and B are **independent** when $P(B \mid A) = P(B)$. The multiplication rule for intersections then becomes $P(A \text{ and } B) = P(A)P(B)$.

A Venn diagram, together with the general addition rule, can be helpful in finding probabilities of the union of two events $P(A \text{ or } B)$ or the joint probability $P(A \text{ and } B)$. The joint probability $P(A \text{ and } B)$ can also be found using the general multiplication rule: $P(A \text{ and } B) = P(A)P(B \mid A) = P(B)P(A \mid B)$.

Constructing a table is a good approach for determining a conditional probability.

In problems with several stages, draw a **tree diagram** to organize use of the multiplication and addition rules.

SECTION 6.3 EXERCISES

6.66 NOBEL PRIZE WINNERS The numbers of Nobel Prize laureates in selected sciences, 1901 to 1998, are shown in the following table by location of award-winning research:[10]

Country	Physics	Chemistry	Physiology/medicine
United States	70	46	82
United Kingdom	21	26	24
Germany	61	17	29
France	25	11	7
Soviet Union	10	7	1
Japan	4	3	1

If a laureate is selected at random, what is the probability that

(a) his or her award was in chemistry?

(b) the award was won by someone from the United States?

(c) the awardee was from the United States, given that the award was for physiology/medicine?

(d) the award was for physiology/medicine, given that the awardee was from the United States?

(e) Interpret each of your results in parts (a) through (d) in terms of percents.

6.67 ACADEMIC DEGREES Here are the counts (in thousands) of earned degrees in the United States in a recent year, classified by level and by the sex of the degree recipient:

	Bachelor's	Master's	Professional	Doctorate	Total
Female	616	194	30	16	856
Male	529	171	44	26	770
Total	1145	365	74	42	1626

(a) If you choose a degree recipient at random, what is the probability that the person you choose is a woman?

(b) What is the conditional probability that you choose a woman, given that the person chosen received a professional degree?

(c) Are the events "choose a woman" and "choose a professional degree recipient" independent? How do you know?

6.68 PICK A CARD The suit of 13 hearts (A, 2 to 10, J, Q, K) from a standard deck of cards is placed in a hat. The cards are thoroughly mixed and a student reaches into the hat and selects two cards without replacement.

(a) What is the probability that the first card selected is the jack?

(b) Given that the first card selected is the jack, what is the probability that the second card is the 5?

(c) What is the probability of selecting the jack on the first draw and then the 5?

(d) What is the probability that both cards selected are greater than 5 (when the ace is considered "low")?

6.69 ACADEMIC DEGREES, II Exercise 6.67 gives the counts (in thousands) of earned degrees in the United States in a recent year. Use these data to answer the following questions.

(a) What is the probability that a randomly chosen degree recipient is a man?

(b) What is the conditional probability that the person chosen received a bachelor's degree, given that he is a man?

(c) Use the multiplication rule to find the joint probability of choosing a male bachelor's degree recipient. Check your result by finding this probability directly from the table of counts.

6.70 TEENAGE DRIVERS An insurance company has the following information about drivers aged 16 to 18 years: 20% are involved in accidents each year; 10% in this age group are A students; among those involved in an accident, 5% are A students.

(a) Let A be the event that a young driver is an A student and C the event that a young driver is involved in an accident this year. State the information given in terms of probabilities and conditional probabilities for the events A and C.

(b) What is the probability that a randomly chosen young driver is an A student and is involved in an accident?

6.71 MORE ON TEENAGE DRIVERS Use your work from Exercise 6.70 to find the percent of A students who are involved in accidents. (Start by expressing this as a conditional probability.)

6.72 Suppose that in Exercise 6.57 (page 370) the treasurer also feels that if the dollar does not fall, there is probability 0.2 that the Japanese supplier will demand that the contract be renegotiated. What is the probability that the supplier will demand renegotiation?

6.73 MULTIPLE-CHOICE EXAM STRATEGIES An examination consists of multiple-choice questions, each having five possible answers. Linda estimates that she has probability 0.75 of knowing the answer to any question that may be asked. If she does not know the answer, she will guess, with conditional probability 1/5 of being correct. What is the probability that Linda gives the correct answer to a question? (Draw a tree diagram to guide the calculation.)

6.74 ELECTION MATH The voters in a large city are 40% white, 40% black, and 20% Hispanic. (Hispanics may be of any race in official statistics, but in this case we are speaking of political blocks.) A black mayoral candidate anticipates attracting 30% of the white vote, 90% of the black vote, and 50% of the Hispanic vote. Draw a tree diagram with probabilities for the race (white, black, or Hispanic) and vote (for or against the candidate) of a randomly chosen voter. What percent of the overall vote does the candidate expect to get?

6.75 In the setting of Exercise 6.73, find the conditional probability that Linda knows the answer, given that she supplies the correct answer. (*Hint:* Use the result of Exercise 6.73 and the definition of conditional probability.)

6.76 GEOMETRIC PROBABILITY Choose a point at random in the square □ with sides $0 \le x \le 1$ and $0 \le y \le 1$. This means that the probability that the point falls in any region within the square is the area of that region. Let X be the x coordinate and Y the y coordinate of the point chosen. Find the conditional probability $P(Y < 1/2 \mid Y > X)$. (*Hint:* Draw a diagram of the square and the events $Y < 1/2$ and $Y > X$.)

6.77 INSPECTING SWITCHES A shipment contains 10,000 switches. Of these, 1000 are bad. An inspector draws switches at random, so that each switch has the same chance to be drawn.

(a) Draw one switch. What is the probability that the switch you draw is bad? What is the probability that it is not bad?

(b) Suppose the first switch drawn is bad. How many switches remain? How many of them are bad? Draw a second switch at random. What is the conditional probability that this switch is bad?

(c) Answer the questions in (b) again, but now suppose that the first switch drawn is not bad.

Comment: Knowing the result of the first trial changes the conditional probability for the second trial, so the trials are not independent. But because the shipment is large, the probabilities change very little. The trials are almost independent.

CHAPTER REVIEW

Probability describes the pattern of chance outcomes. Probability calculations provide the basis for inference. When data are produced by random sampling or randomized comparative experiments, the laws of probability answer the question, "What would happen if we did this very many times?" Probability is used to describe the long-term regularity that results from many repetitions of the same random phenomenon. The reasoning of statistical inference rests on asking "How often would this method give a correct answer if I used it very many times?" This chapter developed a probability model, including rules and tools that will help you describe the behavior of statistics from random samples in later chapters. Here are the most important things you should be able to do after studying this chapter.

PROBABILITY RULES

1. Describe the sample space of a random phenomenon. For a finite number of outcomes, use the multiplication principle to determine the number of outcomes, and use counting techniques, Venn diagrams, and tree diagrams to determine simple probabilities. For the continuous case, use geometric areas to find probabilities (areas under simple density curves) of events (intervals on the horizontal axis).

2. Know the probability rules and be able to apply them to determine probabilities of defined events. In particular, determine if a given assignment of probabilities is valid.

3. Determine if two events are disjoint, complementary, or independent. Find unions and intersections of two or more events.

4. Use Venn diagrams to picture relationships among several events.

5. Use the general addition rule to find probabilities that involve overlapping events.

6. Understand the idea of independence. Judge when it is reasonable to assume independence as part of a probability model.

7. Use the multiplication rule for independent events to find the probability that all of several independent events occur.

8. Use the multiplication rule for independent events in combination with other probability rules to find the probabilities of complex events.

9. Understand the idea of conditional probability. Find conditional probabilities for individuals chosen at random from a table of counts of possible outcomes.

10. Use the general multiplication rule to find the joint probability $P(A \text{ and } B)$ from $P(A)$ and the conditional probability $P(B \mid A)$.

11. Construct tree diagrams to organize the use of the multiplication and addition rules to solve problems with several stages.

CHAPTER 6 REVIEW EXERCISES

6.78 WHO GETS TO GO? Abby, Deborah, Julie, Sam, and Roberto work in a firm's public relations office. Their employer must choose two of them to attend a conference in Paris. To avoid unfairness, the choice will be made by drawing two names from a hat. (This is an SRS of size 2.)

(a) Write down all possible choices of two of the five names. This is the sample space.

(b) The random drawing makes all choices equally likely. What is the probability of each choice?

(c) What is the probability that Julie is chosen?

(d) What is the probability that neither of the two men (Sam and Roberto) is chosen?

6.79 ARE YOU MY (BLOOD) TYPE? All human blood can be "ABO-typed" as one of O, A, B, or AB, but the distribution of the types varies a bit among groups of people. Here is the distribution of blood types for a randomly chosen person in the United States:

Blood type:	O	A	B	AB
U.S. probability:	0.45	0.40	0.11	?

(a) What is the probability of type AB blood in the United States?

(b) An individual with type B blood can safely receive transfusions only from persons with type B or type O blood. What is the probability that the husband of a woman with type B blood is an acceptable blood donor for her?

(c) What is the probability that in a randomly chosen couple the wife has type B blood and the husband has type A?

(d) What is the probability that one of a randomly chosen couple has type A blood and the other has type B?

(e) What is the probability that at least one of a randomly chosen couple has type O blood?

6.80 The distribution of blood types in China differs from the U.S. distribution given in the previous exercise:

Blood type:	O	A	B	AB
China probability:	0.35	0.27	0.26	0.12

Choose an American and a Chinese at random, independently of each other.

(a) What is the probability that both have type O blood?

(b) What is the probability that both have the same blood type?

6.81 INCOME AND SAVINGS A sample survey chooses a sample of households and measures their annual income and their savings. Some events of interest are

A = the household chosen has income at least $100,000

C = the household chosen has at least $50,000 in savings

Based on this sample survey, we estimate that $P(A) = 0.07$ and $P(C) = 0.2$.

(a) We want to find the probability that a household either has income at least $100,000 *or* savings at least $50,000. Explain why we do not have enough information to find this probability. What additional information is needed?

(b) We want to find the probability that a household has income at least $100,000 *and* savings at least $50,000. Explain why we do not have enough information to find this probability. What additional information is needed?

6.82 SCREENING JOB APPLICANTS A company retains a psychologist to assess whether job applicants are suited for assembly-line work. The psychologist classifies applicants as *A* (well suited), *B* (marginal), or *C* (not suited). The company is concerned about event *D*: an employee leaves the company within a year of being hired. Data on all people hired in the past 5 years give these probabilities:

$P(A) = 0.4$	$P(B) = 0.3$	$P(C) = 0.3$
$P(A \text{ and } D) = 0.1$	$P(B \text{ and } D) = 0.1$	$P(C \text{ and } D) = 0.2$

Sketch a Venn diagram of the events *A*, *B*, *C*, and *D* and mark on your diagram the probabilities of all combinations of psychological assessment and leaving (or not) within a year. What is $P(D)$, the probability that an employee leaves within a year?

6.83 SUICIDES Here is a two-way table of suicides committed in a recent year, classified by the gender of the victim and whether or not a firearm was used:

	Male	Female	Total
Firearm	16,381	2,559	18,940
Other	9,034	3,536	12,570
Total	25,415	6,095	31,510

Choose a suicide at random. Find the following probabilities.

(a) $P(\text{a firearm was used})$

(b) $P(\text{firearm} \mid \text{female})$

(c) P(female and firearm)

(d) P(firearm | male)

(e) P(male | firearm)

6.84 **AT THE GYM** Many conditional probability calculations are just common sense made automatic. For example, 10% of adults belong to health clubs, and 40% of these health club members go to the club at least twice a week. What percent of all adults go to a health club at least twice a week? Write the information in terms of probabilities and use the general multiplication rule.

6.85 **TOSS TWO COINS** Independence of events is not always obvious. Toss two balanced coins independently. The four possible combinations of heads and tails in order each have probability 0.25. The events

A = head on the first toss

B = both tosses have the same outcome

may seem intuitively related. Show that $P(B \mid A) = P(B)$, so that A and B are in fact independent.

6.86 **BYPASS SURGERY** John has coronary artery disease. He and his doctor must decide between medical management of the disease and coronary bypass surgery. Because John has been quite active, he is concerned about his quality of life as well as length of life. He wants to make the decision that will maximize the probability of the event A that he survives for 5 years and is able to carry on moderate activity during that time. The doctor makes the following probability estimates for patients of John's age and condition:

- Under medical management, $P(A) = 0.7$.
- There is probability 0.05 that John will not survive bypass surgery, probability 0.10 that he will survive with serious complications, and probability 0.85 that he will survive the surgery without complications.
- If he survives with complications, the conditional probability of the desired outcome A is 0.73. If there are no serious complications, the conditional probability of A is 0.76.

Draw a tree diagram that summarizes this information. Then calculate $P(A)$ assuming that John chooses the surgery. Does surgery or medical management offer him a better chance of achieving his goal?

6.87 **POLL ON SENSITIVE ISSUES** It is difficult to conduct sample surveys on sensitive issues because many people will not answer questions if the answers might embarrass them. "Randomized response" is an effective way to guarantee anonymity while collecting information on topics such as student cheating or sexual behavior. Here is the idea. To ask a sample of students whether they have plagiarized a term paper while in college, have each student toss a coin in private. If the coin lands "heads" *and* they have not plagiarized, they are to answer "No." Otherwise they are to give "Yes" as their answer. Only the student knows whether the answer reflects the truth or just the coin toss, but the researchers can use a proper random sample with follow-up for nonresponse and other good sampling practices.

Suppose that in fact the probability is 0.3 that a randomly chosen student has plagiarized a paper. Draw a tree diagram in which the first stage is tossing the coin and the second is the truth about plagiarism. The outcome at the end of each branch is the answer given to the randomized-response question. What is the probability of a "No" answer in the randomized-response poll? If the probability of plagiarism were 0.2, what would be the probability of a "No" response on the poll? Now suppose that you get 39% "No" answers in a randomized-response poll of a large sample of students at your college. What do you estimate to be the percent of the population who have plagiarized a paper?

NOTES AND DATA SOURCES

1. An informative and entertaining account of the origins of probability theory is Florence N. David, *Games, Gods and Gambling*, Charles Griffin, London, 1962.
2. From the EESEE story "Home-Field Advantage." The study is W. Hurley, "What sort of tournament should the World Series be?" *Chance*, 6, No. 2 (1993), pp. 31–33.
3. You can find a mathematical explanation of Benford's Law in Ted Hill, "The first-digit phenomenon," *American Scientist*, 86 (1996), pp. 358–363, and Ted Hill, "The difficulty of faking data," *Chance*, 12, No. 3 (1999), pp. 27–31. Applications to fraud detection are discussed in the second paper by Hill and in Mark A. Nigrini, "I've got your number," *Journal of Accountancy*, May 1999, available online at www.aicpa.org/pubs/jofa/joaiss.htm.
4. Corey Kilgannon, "When New York is on the end of the line," *New York Times*, November 7, 1999.
5. From the Dupont Automotive North America Color Popularity Survey, reported at www.dupont.com/automotive/.
6. This and similar psychology experiments are reported by A. Tversky and D. Kahneman, "Extensional versus intuitive reasoning: the conjunction fallacy in probability judgement," *Psychological Review*, 90 (1983), pp. 293–315.
7. These probabilities come from studies by the sociologist Harry Edwards, reported in the *New York Times*, February 25, 1986.
8. This example is modeled on Benjamin A. Barnes, "An overview of the treatment of end-stage renal disease and a consideration of some of the consequences," in J. P. Bunker, B. A. Barnes, and F. W. Mosteller (eds.), *Costs, Risks and Benefits of Surgery*, Oxford University Press, New York, 1977, pp. 325–341. The probabilities are recent estimates based on data from the United Network for Organ Sharing (www.unos.org) and Rebecca D. Williams, "Living day-to-day with kidney dialysis," Food and Drug Administration, www.fda.gov.
9. Probabilities from trials with 2897 people known to be free of HIV antibodies and 673 people known to be infected are reported in J. Richard George, "Alternative

specimen sources: methods for confirming positives," 1998 Conference on the Laboratory Science of HIV, found online at the Centers for Disease Control and Prevention, www.cdc.gov.
10. Data from the National Science Foundation, as reported in the *Statistical Abstract of the United States, 2000*.

The Granger Collection, New York

JAKOB BERNOULLI

The Law of Large Numbers

In three generations, the remarkable Bernoulli family of Basel, Switzerland, produced eight mathematicians, several of them outstanding. Five of them, including *Jakob (1654–1705)* and his brother Johann, made significant contributions to the early study of probability. By 1689 Jakob had published his *law of large numbers* in probability theory. The law of large numbers, which we will meet in this chapter, says that if an experiment is repeated many times, then the relative frequency with which an event occurs equals the probability of the event. Although he may be best known for his work in probability theory, Jakob made contributions in other areas as well. He published important work on the connections between logic and algebra, on geometry, and on infinite series.

Jakob Bernoulli's most important work was *Ars Conjectandi* (The Art of Conjecture), published in Basel in 1713, eight years after his death. The book was incomplete at the time of his death but it is still a significant accomplishment in the theory of probability. In the book Bernoulli reviewed the work of others on probability and gave many examples on how much one could expect to win playing various games of chance. He also offered the first proof of the binomial theorem for arbitrary positive integral powers.

Jakob Bernoulli held the chair of mathematics at the University of Basel from 1687 until his death in 1705, when he was succeeded by his brother Johann. Jakob had always been fascinated by the logarithmic spiral, several of whose properties he discovered, and he directed that it be carved on his tombstone with the Latin inscription *Eadem Mutata Resurgo*,"Though changed I shall arise the same."

If an experiment is repeated many times, then the relative frequency with which an event occurs equals the probability of the event.