Chapter

2 PROBABILITY

CHAPTER OUTLINE

2.1 Introduction
2.2 Sample Spaces and the Algebra of Sets
2.3 The Probability Function
2.4 Conditional Probability
2.5 Independence
2.6 Combinatorics
2.7 Combinatorial Probability
2.8 Taking a Second Look at Statistics (Monte Carlo Techniques)

One of the most influential of seventeenth-century mathematicians, Fermat earned his living as a lawyer and administrator in Toulouse. He shares credit with Descartes for the invention of analytic geometry, but his most important work may have been in number theory. Fermat did not write for publication, preferring instead to send letters and papers to friends. His correspondence with Pascal was the starting point for the development of a mathematical theory of probability.

—Pierre de Fermat (1601–1665)

Pascal was the son of a nobleman. A prodigy of sorts, he had already published a treatise on conic sections by the age of sixteen. He also invented one of the early calculating machines to help his father with accounting work. Pascal's contributions to probability were stimulated by his correspondence, in 1654, with Fermat. Later that year he retired to a life of religious meditation.

—Blaise Pascal (1623–1662)

2.1 INTRODUCTION

Experts have estimated that the likelihood of any given UFO sighting being genuine is on the order of one in one hundred thousand. Since the early 1950s, some ten thousand sightings have been reported to civil authorities. What is the probability that at least one of those objects was, in fact, an alien spacecraft? In 1978, Pete Rose of the Cincinnati Reds set a National League record by batting safely in forty-four consecutive games. How unlikely was that event, given that Rose was a lifetime .303 hitter? By definition, the *mean free path* is the average distance a molecule in a gas travels before colliding with another molecule. How likely is it that the distance a molecule travels between collisions will be at least twice its mean free path? Suppose a boy's mother and father both have genetic markers for sickle cell anemia, but neither parent exhibits any of the disease's symptoms. What are the chances that their son will also be asymptomatic? What are the odds that a poker player is dealt a full house or that a craps-shooter makes his "point"? If a woman has lived to age seventy, how likely is it that she will die before her ninetieth birthday? In 1994, Tom Foley was Speaker of the House and running for re-election. The day after the election, his race had still not been "called" by any of the networks: he trailed his Republican challenger by 2174 votes, but 14,000 absentee ballots remained to be counted. Foley, however, conceded. Should he have waited for the absentee ballots to be counted, or was his defeat at that point a virtual certainty?

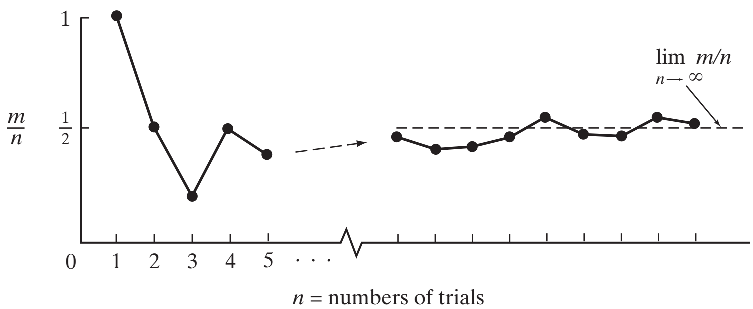
As the nature and variety of these questions would suggest, probability is a subject with an extraordinary range of real-world, everyday applications. What began as an exercise in understanding games of chance has proven to be useful everywhere. Maybe even more remarkable is the fact that the solutions to all of these diverse questions are rooted in just a handful of definitions and theorems. Those results, together with the problem-solving techniques they empower, are the sum and substance of Chapter 2. We begin, though, with a bit of history.

THE EVOLUTION OF THE DEFINITION OF PROBABILITY

Over the years, the definition of probability has undergone several revisions. There is nothing contradictory in the multiple definitions—the changes primarily reflected the need for greater generality and more mathematical rigor. The first formulation (often referred to as the *classical* definition of probability) is credited to Gerolamo Cardano (recall Section 1.3). It applies only to situations where (1) the number of possible outcomes is finite and (2) all outcomes are equally likely. Under those conditions, the probability of an event comprised of m outcomes is the ratio m/n, where n is the total number of (equally likely) outcomes. Tossing a fair, six-sided die, for example, gives $m/n = \frac{3}{6}$ as the probability of rolling an even number (that is, either 2, 4, or 6).

While Cardano's model was well-suited to gambling scenarios (for which it was intended), it was obviously inadequate for more general problems, where outcomes are not equally likely and/or the number of outcomes is not finite. Richard von Mises, a twentieth-century German mathematician, is often credited with avoiding the weaknesses in Cardano's model by defining "empirical" probabilities. In the von Mises approach, we imagine an experiment being repeated over and over again *under presumably identical conditions*. Theoretically, a running tally could be kept of the number of times (m) the outcome belongs to a given event divided by n, the total number of times the experiment is performed. According to von Mises, the probability of the given event is the limit (as n goes to infinity) of the ratio m/n. Figure 2.1.1 illustrates the empirical probability of getting a head by tossing a fair coin: As the number of tosses continues to increase, the ratio m/n converges to $\frac{1}{2}$.

Figure 2.1.1



The von Mises approach definitely shores up some of the inadequacies seen in the Cardano model, but it is not without shortcomings of its own. There is some conceptual inconsistency, for example, in extolling the limit of m/n as a way of defining a probability *empirically*, when the very act of repeating an experiment under identical conditions an infinite number of times is physically impossible. And left unanswered is the question of how large n must be in order for m/n to be a good approximation for lim m/n.

Andrei Kolmogorov, the great Russian probabilist, took a different approach. Aware that many twentieth-century mathematicians were having success developing subjects axiomatically, Kolmogorov wondered whether probability might similarly be defined operationally, rather than as a ratio (like the Cardano model) or as a limit (like the von Mises model). His efforts culminated in a masterpiece of mathematical elegance when he published *Grundbegriffe der Wahrscheinlichkeitsrechnung (Foundations of the Theory of Probability)* in 1933. In essence, Kolmogorov was able to show that a maximum of four simple axioms is necessary and sufficient to define the way any and all probabilities must behave. (These will be our starting point in Section 2.3.)

We begin Chapter 2 with some basic (and, presumably, familiar) definitions from set theory. These are important because probability will eventually be defined as a *set function*—that is, a mapping from a set to a number. Then, with the help of Kolmogorov's axioms in Section 2.3, we will learn how to calculate and manipulate probabilities. The chapter concludes with an introduction to *combinatorics*—the mathematics of systematic counting—and its application to probability.

2.2 Sample Spaces and the Algebra of Sets

The starting point for studying probability is the definition of four key terms: *experiment, sample outcome, sample space*, and *event*. The latter three, all carryovers from classical set theory, give us a familiar mathematical framework within which to work; the former is what provides the conceptual mechanism for casting real-world phenomena into probabilistic terms.

By an *experiment* we will mean any procedure that (1) can be repeated, theoretically, an infinite number of times and (2) has a well-defined set of possible outcomes. Thus, rolling a pair of dice qualifies as an experiment and so does measuring a hypertensive's blood pressure or doing a spectrographic analysis to determine the carbon content of moon rocks. Asking a would-be psychic to draw a picture of an image presumably transmitted by another would-be psychic does *not* qualify as an experiment, because the set of possible outcomes cannot be listed, characterized, or otherwise defined.

Each of the potential eventualities of an experiment is referred to as a *sample outcome*, s, and their totality is called the *sample space*, S. To signify the membership of s in S, we write $s \in S$. Any designated collection of sample outcomes, including individual outcomes, the entire sample space, and the null set, constitutes an *event*. The latter is said to *occur* if the outcome of the experiment is one of the members of the event.

Example 2.2.1

Consider the experiment of flipping a coin three times. What is the sample space? Which sample outcomes make up the event A: Majority of coins show heads?

Think of each sample outcome here as an ordered triple, its components representing the outcomes of the first, second, and third tosses, respectively. Altogether, there are eight different triples, so those eight comprise the sample space:

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

By inspection, we see that four of the sample outcomes in S constitute the event A:

$$A = \{\text{HHH, HHT, HTH, THH}\}$$ ■

Example 2.2.2

Imagine rolling two dice, the first one red, the second one green. Each sample outcome is an ordered pair (face showing on red die, face showing on green die), and the entire sample space can be represented as a 6×6 matrix (see Figure 2.2.1).

Figure 2.2.1

Face showing on red die \ Face showing on green die	1	2	3	4	5	6
1	(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
2	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
3	(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
4	(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
5	(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
6	(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

A

Gamblers are often interested in the event A that the sum of the faces showing is a 7. Notice in Figure 2.2.1 that the sample outcomes contained in A are the six diagonal entries, (1, 6), (2, 5), (3, 4), (4, 3), (5, 2), and (6, 1). ■

Example 2.2.3

A local TV station advertises two newscasting positions. If three women (W_1, W_2, W_3) and two men (M_1, M_2) apply, the "experiment" of hiring two coanchors generates a sample space of ten outcomes:

$$S = \{(W_1, W_2), (W_1, W_3), (W_2, W_3), (W_1, M_1), (W_1, M_2), (W_2, M_1), (W_2, M_2), (W_3, M_1), (W_3, M_2), (M_1, M_2)\}$$

Does it matter here that the two positions being filled are equivalent? Yes. If the station were seeking to hire, say, a sports announcer and a weather forecaster, the number of possible outcomes would be twenty: (W_2, M_1), for example, would represent a different staffing assignment than (M_1, W_2). ■

Example 2.2.4

The number of sample outcomes associated with an experiment need not be finite. Suppose that a coin is tossed until the first tail appears. If the first toss is itself a tail, the outcome of the experiment is T; if the first tail occurs on the second toss, the outcome is HT; and so on. Theoretically, of course, the first tail may *never* occur, and the infinite nature of S is readily apparent:

$$S = \{\text{T}, \text{HT}, \text{HHT}, \text{HHHT}, \ldots\}$$ ■

Example 2.2.5

There are three ways to indicate an experiment's sample space. If the number of possible outcomes is small, we can simply list them, as we did in Examples 2.2.1 through 2.2.3. In some cases it may be possible to *characterize* a sample space by showing the structure its outcomes necessarily possess. This is what we did in Example 2.2.4. A third option is to state a mathematical formula that the sample outcomes must satisfy.

A computer programmer is running a subroutine that solves a general quadratic equation, $ax^2 + bx + c = 0$. Her "experiment" consists of choosing values for the three coefficients a, b, and c. Define (1) S and (2) the event A: Equation has two equal roots.

First, we must determine the sample space. Since presumably no combinations of finite a, b, and c are inadmissible, we can characterize S by writing a series of inequalities:

$$S = \{(a, b, c) : -\infty < a < \infty, -\infty < b < \infty, -\infty < c < \infty\}$$

Defining A requires the well-known result from algebra that a quadratic equation has equal roots if and only if its discriminant, $b^2 - 4ac$, vanishes. Membership in A, then, is contingent on a, b, and c satisfying an equation:

$$A = \{(a, b, c) : b^2 - 4ac = 0\}$$ ■

Questions

2.2.1. A graduating engineer has signed up for three job interviews. She intends to categorize each one as being either a "success" or a "failure" depending on whether it leads to a plant trip. Write out the appropriate sample space. What outcomes are in the event A: Second success occurs on third interview? In B: First success never occurs? (*Hint:* Notice the similarity between this situation and the coin-tossing experiment described in Example 2.2.1.)

2.2.2. Three dice are tossed, one red, one blue, and one green. What outcomes make up the event A that the sum of the three faces showing equals 5?

2.2.3. An urn contains six chips numbered 1 through 6. Three are drawn out. What outcomes are in the event "Second smallest chip is a 3"? Assume that the order of the chips is irrelevant.

2.2.4. Suppose that two cards are dealt from a standard 52-card poker deck. Let A be the event that the sum of the two cards is 8 (assume that aces have a numerical value of 1). How many outcomes are in A?

2.2.5. In the lingo of craps-shooters (where two dice are tossed and the underlying sample space is the matrix pictured in Figure 2.2.1) is the phrase "making a hard eight." What might that mean?

2.2.6. A poker deck consists of fifty-two cards, representing thirteen denominations (2 through ace) and four suits (diamonds, hearts, clubs, and spades). A five-card hand is called a *flush* if all five cards are in the same suit but not all five denominations are consecutive. Pictured below is a flush in hearts. Let N be the set of five cards in hearts that are *not* flushes. How many outcomes are in N? [*Note:* In poker, the denominations (A, 2, 3, 4, 5) are considered to be consecutive (in addition to sequences such as (8, 9, 10, J, Q)).]

		Denominations												
		2	3	4	5	6	7	8	9	10	J	Q	K	A
	D													
Suits	H	X	X				X				X	X		
	C													
	S													

2.2.7. Let P be the set of right triangles with a 5″ hypotenuse and whose height and length are a and b, respectively. Characterize the outcomes in P.

2.2.8. Suppose a baseball player steps to the plate with the intention of trying to "coax" a base on balls by never swinging at a pitch. The umpire, of course, will necessarily call each pitch either a ball (B) or a strike (S). What outcomes make up the event A, that a batter walks on the sixth pitch? (*Note:* A batter "walks" if the fourth ball is called before the third strike.)

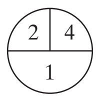
2.2.9. A telemarketer is planning to set up a phone bank to bilk widows with a Ponzi scheme. His past experience (prior to his most recent incarceration) suggests that each phone will be in use half the time. For a given phone at a given time, let 0 indicate that the phone is available and let 1 indicate that a caller is on the line. Suppose that the telemarketer's "bank" is comprised of four telephones.

(a) Write out the outcomes in the sample space.

(b) What outcomes would make up the event that exactly two phones are being used?

(c) Suppose the telemarketer had k phones. How many outcomes would allow for the possibility that at most one more call could be received? (*Hint:* How many lines would have to be busy?)

2.2.10. Two darts are thrown at the following target:



(a) Let (u, v) denote the outcome that the first dart lands in region u and the second dart in region v. List the sample space of (u, v)'s.

(b) List the outcomes in the sample space of *sums*, $u + v$.

2.2.11. A woman has her purse snatched by two teenagers. She is subsequently shown a police lineup consisting of five suspects, including the two perpetrators. What is the sample space associated with the experiment "Woman picks two suspects out of lineup"? Which outcomes are in the event A: She makes at least one incorrect identification?

2.2.12. Consider the experiment of choosing coefficients for the quadratic equation $ax^2 + bx + c = 0$. Characterize the values of a, b, and c associated with the event A: Equation has complex roots.

2.2.13. In the game of craps, the person rolling the dice (the *shooter*) wins outright if his first toss is a 7 or an 11. If his first toss is a 2, 3, or 12, he loses outright. If his first roll is something else, say a 9, that number becomes his "point" and he keeps rolling the dice until he either rolls another 9, in which case he wins, or a 7, in which case he loses. Characterize the sample outcomes contained in the event "Shooter wins with a point of 9."

2.2.14. A probability-minded despot offers a convicted murderer a final chance to gain his release. The prisoner is given twenty chips, ten white and ten black. All twenty are to be placed into two urns, according to any allocation scheme the prisoner wishes, with the one proviso being that each urn contain at least one chip. The executioner will then pick one of the two urns at random and from that urn, one chip at random. If the chip selected is white, the prisoner will be set free; if it is black, he "buys the farm." Characterize the sample space describing the prisoner's possible allocation options. (Intuitively, which allocation affords the prisoner the greatest chance of survival?)

2.2.15. Suppose that ten chips, numbered 1 through 10, are put into an urn at one minute to midnight, and chip number 1 is quickly removed. At one-half minute to midnight, chips numbered 11 through 20 are added to the urn, and chip number 2 is quickly removed. Then at one-fourth minute to midnight, chips numbered 21 to 30 are added to the urn, and chip number 3 is quickly removed. If that procedure for adding chips to the urn continues, how many chips will be in the urn at midnight (157)?

UNIONS, INTERSECTIONS, AND COMPLEMENTS

Associated with events defined on a sample space are several operations collectively referred to as the *algebra of sets*. These are the rules that govern the ways in which one event can be combined with another. Consider, for example, the game of craps described in Question 2.2.13. The shooter wins on his initial roll if he throws either a 7 or an 11. In the language of the algebra of sets, the event "Shooter rolls a 7 *or* an 11" is the *union* of two simpler events, "Shooter rolls a 7" and "Shooter rolls an 11." If E denotes the union and if A and B denote the two events making up the union, we write $E = A \cup B$. The next several definitions and examples illustrate those portions of the algebra of sets that we will find particularly useful in the chapters ahead.

Definition 2.2.1
Let A and B be any two events defined over the same sample space S. Then

a. The *intersection* of A and B, written $A \cap B$, is the event whose outcomes belong to both A and B.

b. The *union* of A and B, written $A \cup B$, is the event whose outcomes belong to either A or B or both.

Example 2.2.6

A single card is drawn from a poker deck. Let A be the event that an ace is selected:

$$A = \{\text{ace of hearts, ace of diamonds, ace of clubs, ace of spades}\}$$

Let B be the event "Heart is drawn":

$$B = \{\text{2 of hearts, 3 of hearts}, \ldots, \text{ace of hearts}\}$$

Then

$$A \cap B = \{\text{ace of hearts}\}$$

and

$$A \cup B = \{\text{2 of hearts, 3 of hearts}, \ldots, \text{ace of hearts, ace of diamonds, ace of clubs, ace of spades}\}$$

(Let C be the event "Club is drawn." Which cards are in $B \cup C$? In $B \cap C$?) ■

Example 2.2.7 Let A be the set of x's for which $x^2 + 2x = 8$; let B be the set for which $x^2 + x = 6$. Find $A \cap B$ and $A \cup B$.

Since the first equation factors into $(x + 4)(x - 2) = 0$, its solution set is $A = \{-4, 2\}$. Similarly, the second equation can be written $(x + 3)(x - 2) = 0$, making $B = \{-3, 2\}$. Therefore,

$$A \cap B = \{2\}$$

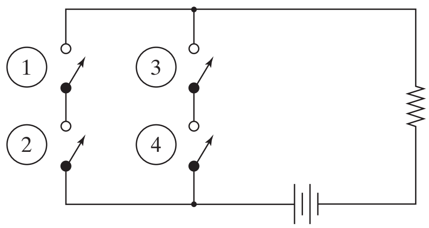
and

$$A \cup B = \{-4, -3, 2\}$$

■

Example 2.2.8 Consider the electrical circuit pictured in Figure 2.2.2. Let A_i denote the event that switch i fails to close, $i = 1, 2, 3, 4$. Let A be the event "Circuit is not completed." Express A in terms of the A_i's.

Figure 2.2.2



Call the ① and ② switches line a; call the ③ and ④ switches line b. By inspection, the circuit fails only if *both* line a and line b fail. But line a fails only if *either* ① *or* ② (or both) fail. That is, the event that line a fails is the union $A_1 \cup A_2$. Similarly, the failure of line b is the union $A_3 \cup A_4$. The event that the circuit fails, then, is an intersection:

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

■

Definition 2.2.2
Events A and B defined over the same sample space are said to be *mutually exclusive* if they have no outcomes in common—that is, if $A \cap B = \emptyset$, where $\emptyset$ is the null set.

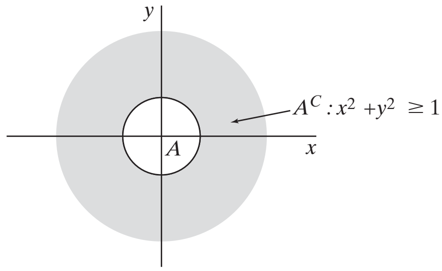
Example 2.2.9 Consider a single throw of two dice. Define A to be the event that the *sum* of the faces showing is odd. Let B be the event that the two faces themselves are odd. Then clearly, the intersection is empty, the sum of two odd numbers necessarily being even. In symbols, $A \cap B = \emptyset$. (Recall the event $B \cap C$ asked for in Example 2.2.6.) ■

Definition 2.2.3
Let A be any event defined on a sample space S. The *complement* of A, written A^C, is the event consisting of all the outcomes in S other than those contained in A.

Example 2.2.10 Let A be the set of (x, y)'s for which $x^2 + y^2 < 1$. Sketch the region in the xy-plane corresponding to A^C.

From analytic geometry, we recognize that $x^2 + y^2 < 1$ describes the interior of a circle of radius 1 centered at the origin. Figure 2.2.3 shows the complement—the points on the circumference of the circle and the points outside the circle.

Figure 2.2.3



The notions of union and intersection can easily be extended to more than two events. For example, the expression $A_1 \cup A_2 \cup \cdots \cup A_k$ defines the set of outcomes belonging to *any* of the A_i's (or to any combination of the A_i's). Similarly, $A_1 \cap A_2 \cap \cdots \cap A_k$ is the set of outcomes belonging to *all* of the A_i's.

Example 2.2.11

Suppose the events $A_1, A_2, \ldots, A_k$ are intervals of real numbers such that

$$A_i = \{x : 0 \leq x < 1/i\}, \quad i = 1, 2, \ldots, k$$

Describe the sets $A_1 \cup A_2 \cup \cdots \cup A_k = \cup_{i=1}^{k} A_i$ and $A_1 \cap A_2 \cap \cdots \cap A_k = \cap_{i=1}^{k} A_i$.

Notice that the A_i's are telescoping sets. That is, A_1 is the interval $0 \leq x < 1$, A_2 is the interval $0 \leq x < \frac{1}{2}$, and so on. It follows, then, that the *union* of the k A_i's is simply A_1 while the *intersection* of the A_i's (that is, their overlap) is A_k.

Questions

2.2.16. Sketch the regions in the xy-plane corresponding to $A \cup B$ and $A \cap B$ if

$$A = \{(x, y): 0 < x < 3, 0 < y < 3\}$$

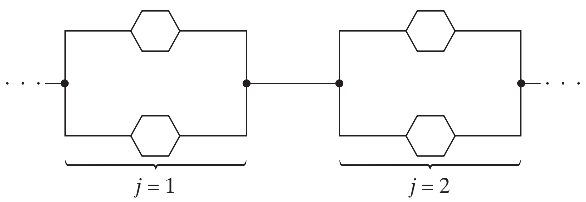
and

$$B = \{(x, y): 2 < x < 4, 2 < y < 4\}$$

2.2.17. Referring to Example 2.2.7, find $A \cap B$ and $A \cup B$ if the two equations were replaced by inequalities: $x^2 + 2x \leq 8$ and $x^2 + x \leq 6$.

2.2.18. Find $A \cap B \cap C$ if $A = \{x: 0 \leq x \leq 4\}$, $B = \{x: 2 \leq x \leq 6\}$, and $C = \{x: x = 0, 1, 2, \ldots\}$.

2.2.19. An electronic system has four components divided into two pairs. The two components of each pair are wired in parallel; the two pairs are wired in series. Let A_{ij} denote the event "ith component in jth pair fails," $i = 1, 2$; $j = 1, 2$. Let A be the event "System fails." Write A in terms of the A_{ij}'s.



2.2.20. Define $A = \{x : 0 \leq x \leq 1\}$, $B = \{x : 0 \leq x \leq 3\}$, and $C = \{x : -1 \leq x \leq 2\}$. Draw diagrams showing each of the following sets of points:

(a) $A^C \cap B \cap C$
(b) $A^C \cup (B \cap C)$
(c) $A \cap B \cap C^C$
(d) $[(A \cup B) \cap C^C]^C$

2.2.21. Let A be the set of five-card hands dealt from a fifty-two-card poker deck, where the denominations of the five cards are all consecutive—for example, (7 of hearts, 8 of spades, 9 of spades, 10 of hearts, jack of diamonds). Let B be the set of five-card hands where the suits of the five cards are all the same. How many outcomes are in the event $A \cap B$?

2.2.22. Suppose that each of the twelve letters in the word

T E S S E L L A T I O N

is written on a chip. Define the events F, R, and C as follows:

F: letters in first half of alphabet
R: letters that are repeated
V: letters that are vowels

Which chips make up the following events?

(a) $F \cap R \cap V$
(b) $F^C \cap R \cap V^C$
(c) $F \cap R^C \cap V$

2.2.23. Let A, B, and C be any three events defined on a sample space S. Show that

(a) the outcomes in $A \cup (B \cap C)$ are the same as the outcomes in $(A \cup B) \cap (A \cup C)$.

(b) the outcomes in $A \cap (B \cup C)$ are the same as the outcomes in $(A \cap B) \cup (A \cap C)$.

2.2.24. Let $A_1, A_2, \ldots, A_k$ be any set of events defined on a sample space S. What outcomes belong to the event

$$(A_1 \cup A_2 \cup \cdots \cup A_k) \cup (A_1^C \cap A_2^C \cap \cdots \cap A_k^C)$$

2.2.25. Let A, B, and C be any three events defined on a sample space S. Show that the operations of union and intersection are *associative* by proving that

(a) $A \cup (B \cup C) = (A \cup B) \cup C = A \cup B \cup C$

(b) $A \cap (B \cap C) = (A \cap B) \cap C = A \cap B \cap C$

2.2.26. Suppose that three events—A, B, and C—are defined on a sample space S. Use the union, intersection, and complement operations to represent each of the following events:

(a) none of the three events occurs

(b) all three of the events occur

(c) only event A occurs

(d) exactly one event occurs

(e) exactly two events occur

2.2.27. What must be true of events A and B if

(a) $A \cup B = B$

(b) $A \cap B = A$

2.2.28. Let events A and B and sample space S be defined as the following intervals:

$$S = \{x : 0 \le x \le 10\}$$
$$A = \{x : 0 < x < 5\}$$
$$B = \{x : 3 \le x \le 7\}$$

Characterize the following events:

(a) A^C

(b) $A \cap B$

(c) $A \cup B$

(d) $A \cap B^C$

(e) $A^C \cup B$

(f) $A^C \cap B^C$

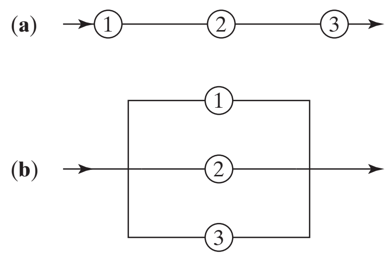
2.2.29. A coin is tossed four times and the resulting sequence of heads and/or tails is recorded. Define the events A, B, and C as follows:

A: exactly two heads appear
B: heads and tails alternate
C: first two tosses are heads

(a) Which events, if any, are mutually exclusive?

(b) Which events, if any, are subsets of other sets?

2.2.30. Pictured below are two organizational charts describing the way upper management vets new proposals. For both models, three vice presidents—1, 2, and 3—each voice an opinion.



For (a), all three must concur if the proposal is to pass; if any one of the three favors the proposal in (b), it passes. Let A_i denote the event that vice president i favors the proposal, $i = 1, 2, 3$, and let A denote the event that the proposal passes. Express A in terms of the A_i's for the two office protocols. Under what sorts of situations might one system be preferable to the other?

EXPRESSING EVENTS GRAPHICALLY: VENN DIAGRAMS

Relationships based on two or more events can sometimes be difficult to express using only equations or verbal descriptions. An alternative approach that can be highly effective is to represent the underlying events graphically in a format known as a *Venn diagram*. Figure 2.2.4 shows Venn diagrams for an intersection, a union, a complement, and two events that are mutually exclusive. In each case, the shaded interior of a region corresponds to the desired event.

Example 2.2.12

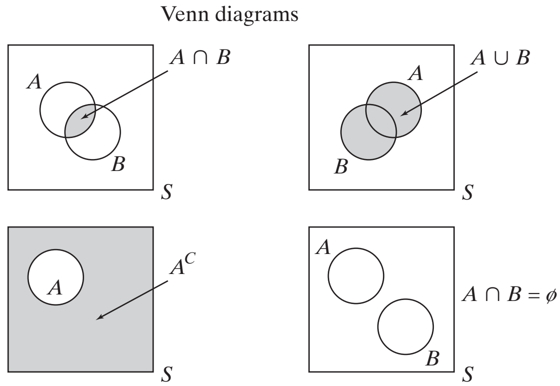
When two events A and B are defined on a sample space, we will frequently need to consider

a. the event that *exactly one* (of the two) occurs.

b. the event that *at most one* (of the two) occurs.

Getting expressions for each of these is easy if we visualize the corresponding Venn diagrams.

Figure 2.2.4



The shaded area in Figure 2.2.5 represents the event E that either A or B, but not both, occurs (that is, *exactly one* occurs).

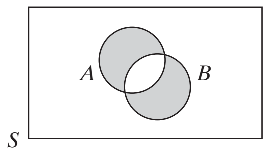


Figure 2.2.5

Just by looking at the diagram we can formulate an expression for E. The portion of A, for example, included in E is $A \cap B^C$. Similarly, the portion of B included in E is $B \cap A^C$. It follows that E can be written as a union:

$$E = (A \cap B^C) \cup (B \cap A^C)$$

(Convince yourself that an equivalent expression for E is $(A \cap B)^C \cap (A \cup B)$.)

Figure 2.2.6 shows the event F that *at most one* (of the two events) occurs. Since the latter includes every outcome except those belonging to *both* A and B, we can write

$$F = (A \cap B)^C$$

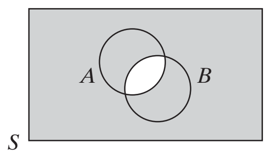


Figure 2.2.6

■

The final example in this section shows two ways of "verifying" identities involving events. The first is nonrigorous and uses Venn diagrams; the second is a formal approach in which every outcome in the left-hand side of the presumed identity is shown to belong to the right-hand side, and vice versa. The particular identity being established is a very useful distributive property for intersections.

Example 2.2.13

Let A, B, and C be any three events defined over the same sample space S. Show that

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \tag{2.2.1}$$

Figure 2.2.7 pictures the set formed by intersecting A with the union of B and C. Similarly, Figure 2.2.8 shows the union of the intersection of A and B with the intersection of A and C. It would appear at this point that the identity is true: The rightmost diagrams in Figures 2.2.7 and 2.2.8 *do* show the same shaded region. Still, this is not a proof. How the events were initially drawn may not do justice to the problem *in general*. Can we be certain, for example, that Equation 2.2.1 remains true if C, for example, is mutually exclusive of A or B? Or if B is a proper subset of A?

Figure 2.2.7

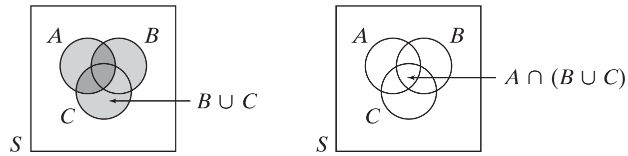
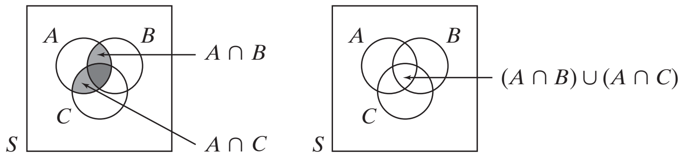


Figure 2.2.8



For a rigorous solution, we need to take the algebraic approach of showing that

(1) $A \cap (B \cup C)$ is contained in $(A \cap B) \cup (A \cap C)$

and

(2) $(A \cap B) \cup (A \cap C)$ is contained in $A \cap (B \cup C)$.

To that end, let $s \in A \cap (B \cup C)$. Then, $s \in A$ and $s \in B \cup C$. But if $s \in B \cup C$, then either $s \in B$ or $s \in C$. If $s \in B$, then $s \in A \cap B$ and $s \in (A \cap B) \cup (A \cap C)$. Likewise, if $s \in C$, it follows that $s \in (A \cap B) \cup (A \cap C)$. Therefore, every sample outcome in $A \cap (B \cup C)$ is also contained in $(A \cap B) \cup (A \cap C)$. Going the other way, assume that $s \in (A \cap B) \cup (A \cap C)$. Therefore, either $s \in (A \cap B)$ or $s \in (A \cap C)$ (or both). Suppose $s \in A \cap B$. Then $s \in A$ and $s \in B$, in which case $s \in A \cap (B \cup C)$. The same conclusion holds if $s \in (A \cap C)$. Thus, every sample outcome in $(A \cap B) \cup (A \cap C)$ is in $A \cap (B \cup C)$. It follows that $A \cap (B \cup C)$ and $(A \cap B) \cup (A \cap C)$ are identical. ■

Questions

2.2.31. During orientation week, the latest Spiderman movie was shown twice at State University. Among the entering class of 6000 freshmen, 850 went to see it the first time, 690 the second time, while 4700 failed to see it either time. How many saw it twice?

2.2.32. Let A and B be any two events. Use Venn diagrams to show that

(a) the complement of their intersection is the union of their complements:

$$(A \cap B)^C = A^C \cup B^C$$

(b) the complement of their union is the intersection of their complements:

$$(A \cup B)^C = A^C \cap B^C$$

(These two results are known as *DeMorgan's laws*.)

2.2.33. Let A, B, and C be any three events. Use Venn diagrams to show that

(a) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

(b) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

2.2.34. Let A, B, and C be any three events. Use Venn diagrams to show that

(a) $A \cup (B \cup C) = (A \cup B) \cup C$

(b) $A \cap (B \cap C) = (A \cap B) \cap C$

2.2.35. Let A and B be any two events defined on a sample space S. Which of the following sets are necessarily subsets of which other sets?

$$A \qquad B \qquad A \cup B \qquad A \cap B \qquad A^C \cap B$$
$$A \cap B^C \qquad (A^C \cup B^C)^C$$

2.2.36. Use Venn diagrams to suggest an equivalent way of representing the following events:

(a) $(A \cap B^C)^C$

(b) $B \cup (A \cup B)^C$

(c) $A \cap (A \cap B)^C$

2.2.37. A total of twelve hundred graduates of State Tech have gotten into medical school in the past several years. Of that number, one thousand earned scores of twenty-seven or higher on the MCAT and four hundred had GPAs that were 3.5 or higher. Moreover, three hundred had MCATs that were twenty-seven or higher *and* GPAs that were 3.5 or higher. What proportion of those twelve hundred graduates got into medical school with an MCAT lower than twenty-seven and a GPA below 3.5?

2.2.38. Let A, B, and C be any three events defined on a sample space S. Let $N(A)$, $N(B)$, $N(C)$, $N(A \cap B)$, $N(A \cap C)$, $N(B \cap C)$, and $N(A \cap B \cap C)$ denote the numbers of outcomes in all the different intersections in which A, B, and C are involved. Use a Venn diagram to suggest a formula for $N(A \cup B \cup C)$. [*Hint:* Start with the sum $N(A) + N(B) + N(C)$ and use the Venn diagram to identify the "adjustments" that need to be made to that sum before it can equal $N(A \cup B \cup C)$.] As a precedent, note that $N(A \cup B) = N(A) + N(B) - N(A \cap B)$. There, in the case of *two* events, subtracting $N(A \cap B)$ is the "adjustment."

2.2.39. A poll conducted by a potential presidential candidate asked two questions: (1) Do you support the candidate's position on taxes? and (2) Do you support the candidate's position on homeland security? A total of twelve hundred responses were received; six hundred said "yes" to the first question and four hundred said "yes" to the second. If three hundred respondents said "no" to the taxes question and "yes" to the homeland security question, how many said "yes" to the taxes question but "no" to the homeland security question?

2.2.40. For two events A and B defined on a sample space S, $N(A \cap B^C) = 15$, $N(A^C \cap B) = 50$, and $N(A \cap B) = 2$. Given that $N(S) = 120$, how many outcomes belong to neither A nor B?

2.3 The Probability Function

Having introduced in Section 2.2 the twin concepts of "experiment" and "sample space," we are now ready to pursue in a formal way the all-important problem of assigning a *probability* to an experiment's outcome—and, more generally, to an event. Specifically, if A is any event defined on a sample space S, the symbol $P(A)$ will denote the *probability of* A, and we will refer to P as the *probability function*. It is, in effect, a mapping from a set (i.e., an event) to a number. The backdrop for our discussion will be the unions, intersections, and complements of set theory; the starting point will be the axioms referred to in Section 2.1 that were originally set forth by Kolmogorov.

If S has a finite number of members, Kolmogorov showed that as few as three axioms are necessary and sufficient for characterizing the probability function P:

Axiom 1. *Let A be any event defined over S. Then $P(A) \geq 0$.*

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

Axiom 3. *Let A and B be any two mutually exclusive events defined over S. Then*

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

When S has an infinite number of members, a fourth axiom is needed:

Axiom 4. *Let $A_1, A_2, \ldots,$ be events defined over S. If $A_i \cap A_j = \emptyset$ for each $i \neq j$, then*

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

From these simple statements come the general rules for manipulating probability functions—rules that apply no matter what specific mathematical forms the functions may happen to take.

SOME BASIC PROPERTIES OF P

Some of the immediate consequences of Kolmogorov's axioms are the results given in Theorems 2.3.1 through 2.3.6. Despite their simplicity, these properties prove to be extraordinarily useful in solving all sorts of problems.

Theorem 2.3.1 $P(A^C) = 1 - P(A)$.

Proof By Axiom 2 and Definition 2.2.3,

$$P(S) = 1 = P(A \cup A^C)$$

But A and A^C are mutually exclusive, so

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

and the result follows.

Theorem 2.3.2 $P(\emptyset) = 0$.

Proof Since $\emptyset = S^C$, $P(\emptyset) = P(S^C) = 1 - P(S) = 0$.

Theorem 2.3.3 *If $A \subset B$, then $P(A) \le P(B)$.*

Proof Note that the event B may be written in the form

$$B = A \cup (B \cap A^C)$$

where A and $(B \cap A^C)$ are mutually exclusive. Therefore,

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

which implies that $P(B) \ge P(A)$ since $P(B \cap A^C) \ge 0$.

Theorem 2.3.4 *For any event A, $P(A) \le 1$.*

Proof The proof follows immediately from Theorem 2.3.3 because $A \subset S$ and $P(S) = 1$.

Theorem 2.3.5 *Let $A_1, A_2, \ldots, A_n$ be events defined over S. If $A_i \cap A_j = \emptyset$ for $i \ne j$, then*

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

Proof The proof is a straightforward induction argument with Axiom 3 being the starting point.

Theorem 2.3.6 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Proof The Venn diagram for $A \cup B$ certainly suggests that the statement of the theorem is true (recall Figure 2.2.4). More formally, we have from Axiom 3 that

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

and

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

Adding these two equations gives

$$P(A) + P(B) = [P(A \cap B^C) + P(B \cap A^C) + P(A \cap B)] + P(A \cap B)$$

By Theorem 2.3.5, the sum in the brackets is $P(A \cup B)$. If we subtract $P(A \cap B)$ from both sides of the equation, the result follows.

The next result is a generalization of Theorem 2.3.6 that considers the probability of the union of n events. We have elected to retain the two-event case, $P(A \cup B)$, as a separate theorem simply for pedagogical reasons.

Theorem 2.3.7 *Let $A_1, A_2, \ldots A_n$ be any n events defined on S. Then*

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

Proof The proof of Theorem 2.3.7 is basically an exercise in bookkeeping. By definition, $A_1 \cup A_2 \cup \cdots \cup A_n$ is the set of outcomes belonging to any of the A_i's individually or to any intersections of the A_i's. The right-hand side of Theorem 2.3.7 is a counting scheme that alternately includes and excludes subsets of outcomes in such a way that each outcome in $A_1 \cup A_2 \cup \cdots \cup A_n$ is included *once and only once* in the calculation of $P(A_1 \cup A_2 \cup \cdots \cup A_n)$.

Consider, for example, the union of *three* events, A_1, A_2, and A_3 as shown in Figure 2.3.1.

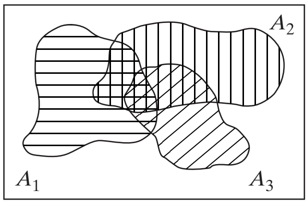


Figure 2.3.1

Notice that simply adding $P(A_1)$, $P(A_2)$, and $P(A_3)$ results in adding $P(A_1 \cap A_2)$ twice, $P(A_1 \cap A_3)$ twice, and $P(A_2 \cap A_3)$ twice. It also results in $P(A_1 \cap A_2 \cap A_3)$ being added three times. The obvious "correction" is to subtract the probabilities associated with $A_1 \cap A_2$, $A_1 \cap A_3$, and $A_2 \cap A_3$ and to *add back* the probability associated with $A_1 \cap A_2 \cap A_3$ since those latter outcomes have previously been added three times and subtracted three times. Therefore,

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= \sum_{i=1}^{3} P(A_i) - \sum_{i<j} P\left(A_i \cap A_j\right) + (-1)^{3+1} P(A_1 \cap A_2 \cap A_3) \\ &= P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) \\ &\quad - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3). \end{aligned}$$

A formal proof of the theorem requires a knowledge of combinatorics and will be deferred until Section 2.6.

Example 2.3.1 Let A and B be two events defined on a sample space S such that $P(A) = 0.3$, $P(B) = 0.5$, and $P(A \cup B) = 0.7$. Find (a) $P(A \cap B)$, (b) $P(A^C \cup B^C)$, and (c) $P(A^C \cap B)$.

a. Transposing the terms in Theorem 2.3.6 yields a general formula for the probability of an intersection:

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

Here

$$P(A \cap B) = 0.3 + 0.5 - 0.7$$
$$= 0.1$$

b. The two cross-hatched regions in Figure 2.3.2 correspond to A^C and B^C. The union of A^C and B^C consists of those regions that have cross-hatching in either or both directions. By inspection, the only portion of S *not* included in $A^C \cup B^C$ is the intersection, $A \cap B$. By Theorem 2.3.1, then

$$P(A^C \cup B^C) = 1 - P(A \cap B)$$
$$= 1 - 0.1$$
$$= 0.9$$

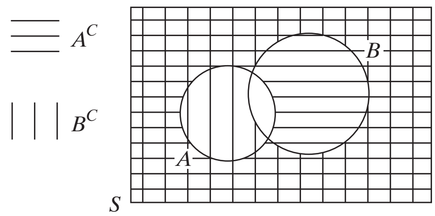


Figure 2.3.2

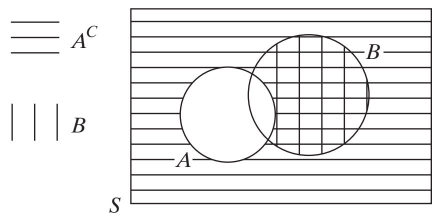


Figure 2.3.3

c. The event $A^C \cap B$ corresponds to the region in Figure 2.3.3 where the cross-hatching extends in *both* directions—that is, everywhere in B except the intersection with A. Therefore,

$$P(A^C \cap B) = P(B) - P(A \cap B)$$
$$= 0.5 - 0.1$$
$$= 0.4$$ ■

Example 2.3.2 Show that

$$P(A \cap B) \geq 1 - P(A^C) - P(B^C)$$

for any two events A and B defined on a sample space S.

From Example 2.3.1a and Theorem 2.3.1,

$$P(A \cap B) = P(A) + P(B) - P(A \cup B)$$
$$= 1 - P(A^C) + 1 - P(B^C) - P(A \cup B)$$

But $P(A \cup B) \leq 1$ from Theorem 2.3.4, so

$$P(A \cap B) \geq 1 - P(A^C) - P(B^C)$$ ■

Example 2.3.3 Two cards are drawn from a poker deck without replacement. What is the probability that the second is higher in rank than the first?

Let A_1, A_2, and A_3 be the events "First card is lower in rank," "First card is higher in rank," and "Both cards have same rank," respectively. Clearly, the three A_i's are mutually exclusive and they account for all possible outcomes, so from Theorem 2.3.5,

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) = P(S) = 1$$

Once the first card is drawn, there are three choices for the second that would have the same rank—that is, $P(A_3) = \frac{3}{51}$. Moreover, symmetry demands that $P(A_1) = P(A_2)$, so

$$2P(A_2) + \frac{3}{51} = 1$$

implying that $P(A_2) = \frac{8}{17}$. ■

Example 2.3.4 Having endured (and survived) the mental trauma that comes from taking two years of chemistry, a year of physics, and a year of biology, Biff decides to test the medical school waters and sends his MCATs to two colleges, X and Y. Based on how his friends have fared, he estimates that his probability of being accepted at X is 0.7, and at Y it is 0.4. He also suspects there is a 75% chance that at least one of his applications will be rejected. What is the probability that he gets at least one acceptance?

Let A be the event "School X accepts him" and B the event "School Y accepts him." We are given that $P(A) = 0.7$, $P(B) = 0.4$, and $P(A^C \cup B^C) = 0.75$. The question is asking for $P(A \cup B)$.

From Theorem 2.3.6,

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

Recall from Question 2.2.32 that $A^C \cup B^C = (A \cap B)^C$, so

$$P(A \cap B) = 1 - P[(A \cap B)^C] = 1 - 0.75 = 0.25$$

It follows that Biff's prospects are not all that bleak—he has an 85% chance of getting in somewhere:

$$\begin{aligned} P(A \cup B) &= 0.7 + 0.4 - 0.25 \\ &= 0.85 \end{aligned}$$

■

Comment Notice that $P(A \cup B)$ varies directly with $P(A^C \cup B^C)$:

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - [1 - P(A^C \cup B^C)] \\ &= P(A) + P(B) - 1 + P(A^C \cup B^C) \end{aligned}$$

If $P(A)$ and $P(B)$, then, are fixed, we get the curious result that Biff's chances of getting at least one acceptance increase if his chances of at least one rejection increase.

Questions

2.3.1. According to a family-oriented lobbying group, there is too much crude language and violence on television. Forty-two percent of the programs they screened had language they found offensive, 27% were too violent, and 10% were considered excessive in both language and violence. What percentage of programs did comply with the group's standards?

2.3.2. Let A and B be any two events defined on S. Suppose that $P(A) = 0.4$, $P(B) = 0.5$, and $P(A \cap B) = 0.1$. What is the probability that A or B but not both occur?

2.3.3. Express the following probabilities in terms of $P(A)$, $P(B)$, and $P(A \cap B)$.

(a) $P(A^C \cup B^C)$

(b) $P(A^C \cap (A \cup B))$

2.3.4. Let A and B be two events defined on S. If the probability that at least one of them occurs is 0.3 and the probability that A occurs but B does not occur is 0.1, what is $P(B)$?

2.3.5. Suppose that three fair dice are tossed. Let A_i be the event that a 6 shows on the ith die, $i = 1, 2, 3$. Does $P(A_1 \cup A_2 \cup A_3) = \frac{1}{2}$? Explain.

2.3.6. Events A and B are defined on a sample space S such that $P((A \cup B)^C) = 0.5$ and $P(A \cap B) = 0.2$. What is the probability that either A or B but not both will occur?

2.3.7. Let $A_1, A_2, \ldots, A_n$ be a series of events for which $A_i \cap A_j = \emptyset$ if $i \neq j$ and $A_1 \cup A_2 \cup \cdots \cup A_n = S$. Let B be any event defined on S. Express B as a union of intersections.

2.3.8. Draw the Venn diagrams that would correspond to the equations (a) $P(A \cap B) = P(B)$ and (b) $P(A \cup B) = P(B)$.

2.3.9. In the game of "odd man out" each player tosses a fair coin. If all the coins turn up the same except for one, the player tossing the different coin is declared the odd man out and is eliminated from the contest. Suppose that three people are playing. What is the probability that someone will be eliminated on the first toss? (*Hint:* Use Theorem 2.3.1.)

2.3.10. An urn contains twenty-four chips, numbered 1 through 24. One is drawn at random. Let A be the event that the number is divisible by 2 and let B be the event that the number is divisible by 3. Find $P(A \cup B)$.

2.3.11. If State's football team has a 10% chance of winning Saturday's game, a 30% chance of winning two weeks from now, and a 65% chance of losing both games, what are their chances of winning exactly once?

2.3.12. Events A_1 and A_2 are such that $A_1 \cup A_2 = S$ and $A_1 \cap A_2 = \emptyset$. Find p_2 if $P(A_1) = p_1$, $P(A_2) = p_2$, and $3p_1 - p_2 = \frac{1}{2}$.

2.3.13. Consolidated Industries has come under considerable pressure to eliminate its seemingly discriminatory hiring practices. Company officials have agreed that during the next five years, 60% of their new employees will be females and 30% will be minorities. One out of four new employees, though, will be a white male. What percentage of their new hires will be minority females?

2.3.14. Three events—A, B, and C—are defined on a sample space, S. Given that $P(A) = 0.2$, $P(B) = 0.1$, and $P(C) = 0.3$, what is the smallest possible value for $P[(A \cup B \cup C)^C]$?

2.3.15. A coin is to be tossed four times. Define events X and Y such that

X: first and last coins have opposite faces
Y: exactly two heads appear

Assume that each of the sixteen head/tail sequences has the same probability. Evaluate

(a) $P(X^C \cap Y)$

(b) $P(X \cap Y^C)$

2.3.16. Two dice are tossed. Assume that each possible outcome has a $\frac{1}{36}$ probability. Let A be the event that the sum of the faces showing is 6, and let B be the event that the face showing on one die is twice the face showing on the other. Calculate $P(A \cap B^C)$.

2.3.17. Let A, B, and C be three events defined on a sample space, S. Arrange the probabilities of the following events from smallest to largest:

(a) $A \cup B$

(b) $A \cap B$

(c) A

(d) S

(e) $(A \cap B) \cup (A \cap C)$

2.3.18. Lucy is currently running two dot-com scams out of a bogus chatroom. She estimates that the chances of the first one leading to her arrest are one in ten; the "risk" associated with the second is more on the order of one in thirty. She considers the likelihood that she gets busted for both to be 0.0025. What are Lucy's chances of avoiding incarceration?

2.4 Conditional Probability

In Section 2.3, we calculated probabilities of certain events by manipulating other probabilities whose values we were given. Knowing $P(A)$, $P(B)$, and $P(A \cap B)$, for example, allows us to calculate $P(A \cup B)$ (recall Theorem 2.3.6). For many real-world situations, though, the "given" in a probability problem goes beyond simply knowing a set of other probabilities. Sometimes, we know *for a fact* that certain events *have already occurred*, and those occurrences may have a bearing on the probability we are trying to find. In short, the probability of an event A may have to be "adjusted" if we know for certain that some related event B has already occurred. Any probability that is revised to take into account the (known) occurrence of other events is said to be a *conditional probability*.

Consider a fair die being tossed, with A defined as the event "6 appears." Clearly, $P(A) = \frac{1}{6}$. But suppose that the die has already been tossed—by someone who refuses to tell us whether or not A occurred but does enlighten us to the extent of confirming that B occurred, where B is the event "Even number appears." What are the chances of A now? Here, common sense can help us: There are three equally likely even numbers making up the event B—one of which satisfies the event A, so the "updated" probability is $\frac{1}{3}$.

Notice that the effect of additional information, such as the knowledge that B has occurred, is to revise—indeed, to *shrink*—the original sample space S to a new set of outcomes S'. In this example, the original S contained six outcomes, the conditional sample space, three (see Figure 2.4.1).

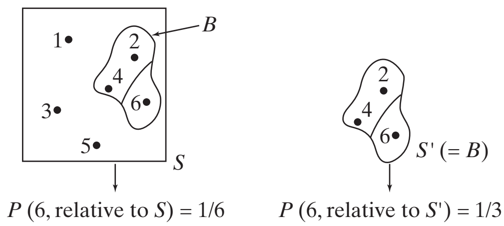


Figure 2.4.1

The symbol $P(A|B)$—read "the probability of A given B"—is used to denote a conditional probability. Specifically, $P(A|B)$ refers to the probability that A *will occur* given that B *has already occurred*.

It will be convenient to have a formula for $P(A|B)$ that can be evaluated in terms of the original S, rather than the revised S'. Suppose that S is a finite sample space with n outcomes, all equally likely. Assume that A and B are two events containing a and b outcomes, respectively, and let c denote the number of outcomes in the intersection of A and B (see Figure 2.4.2). Based on the argument suggested in Figure 2.4.1, the *conditional probability of A given B is the ratio of c to b*. But c/b can be written as the quotient of two other ratios,

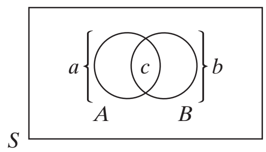


Figure 2.4.2

$$\frac{c}{b} = \frac{c/n}{b/n}$$

so, for this particular case,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{2.4.1}$$

The same underlying reasoning that leads to Equation 2.4.1, though, holds true even when the outcomes are not equally likely or when S is uncountably infinite.

Definition 2.4.1
Let A and B be any two events defined on S such that $P(B) > 0$. The conditional probability of A, assuming that B has already occurred, is written $P(A|B)$ and is given by

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

Comment Definition 2.4.1 can be cross multiplied to give a frequently useful expression for the probability of an intersection. If $P(A|B) = P(A \cap B)/P(B)$, then

$$P(A \cap B) = P(A|B)P(B) \tag{2.4.2}$$

Example 2.4.1

A card is drawn from a poker deck. What is the probability that the card is a club, given that the card is a king?

Intuitively, the answer is $\frac{1}{4}$: The king is equally likely to be a heart, diamond, club, or spade. More formally, let C be the event "Card is a club"; let K be the event "Card is a king." By Definition 2.4.1,

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

But $P(K) = \frac{4}{52}$ and $P(C \cap K) = P(\text{Card is a king of clubs}) = \frac{1}{52}$. Therefore, confirming our intuition,

$$P(C|K) = \frac{1/52}{4/52} = \frac{1}{4}$$

[Notice in this example that the conditional probability $P(C|K)$ is numerically the same as the unconditional probability $P(C)$—they both equal $\frac{1}{4}$. This means that our knowledge that K has occurred gives us no additional insight about the chances of C occurring. Two events having this property are said to be *independent*. We will examine the notion of independence and its consequences in detail in Section 2.5.] ■

Example 2.4.2

Our intuitions can often be fooled by probability problems, even ones that appear to be simple and straightforward. The "two boys" problem described here is an often-cited case in point.

Consider the set of families having two children. Assume that the four possible birth sequences—(younger child is a boy, older child is a boy), (younger child is a boy, older child is a girl), and so on—are equally likely. What is the probability that both children are boys given that at least one is a boy?

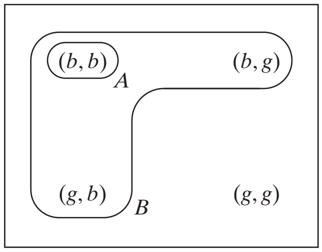
The answer is *not* $\frac{1}{2}$. The correct answer can be deduced from Definition 2.4.1. By assumption, each of the four possible birth sequences—(b, b), (b, g), (g, b), and (g, g)—has a $\frac{1}{4}$ probability of occurring. Let A be the event that both children are boys, and let B be the event that at least one child is a boy. Then

$$P(A|B) = P(A \cap B)/P(B) = P(A)/P(B)$$

since A is a subset of B (so the overlap between A and B is just A). But A has one outcome $\{(b, b)\}$ and B has three outcomes $\{(b, g), (g, b), (b, b)\}$. Applying Definition 2.4.1, then, gives

$$P(A|B) = (1/4)/(3/4) = \frac{1}{3}$$

Another correct approach is to go back to the sample space and deduce the value of $P(A|B)$ from first principles. Figure 2.4.3 shows events A and B defined on the four family types that comprise the sample space S. Knowing that B has occurred redefines the sample space to include *three* outcomes, each now having a $\frac{1}{3}$ probability. Of those three possible outcomes, one—namely, (b, b)—satisfies the event A. It follows that $P(A|B) = \frac{1}{3}$.



S = sample space of two-child families
[outcomes written as (first born, second born)]

Figure 2.4.3 ■

Example 2.4.3 Two events A and B are defined such that (1) the probability that A occurs but B does not occur is 0.2, (2) the probability that B occurs but A does not occur is 0.1, and (3) the probability that neither occurs is 0.6. What is $P(A|B)$?

The three events whose probabilities are given are indicated on the Venn diagram shown in Figure 2.4.4. Since

$$P(\text{Neither occurs}) = 0.6 = P((A \cup B)^C)$$

it follows that

$$P(A \cup B) = 1 - 0.6 = 0.4 = P(A \cap B^C) + P(A \cap B) + P(B \cap A^C)$$

so

$$\begin{aligned} P(A \cap B) &= 0.4 - 0.2 - 0.1 \\ &= 0.1 \end{aligned}$$

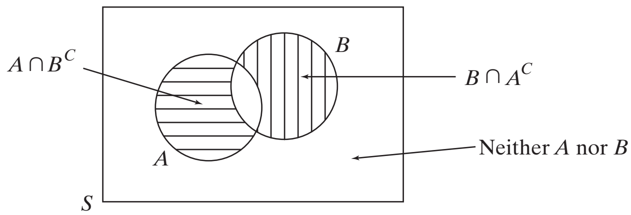


Figure 2.4.4

From Definition 2.4.1, then,

$$\begin{aligned} P(A|B) = \frac{P(A \cap B)}{P(B)} &= \frac{P(A \cap B)}{P(A \cap B) + P(B \cap A^C)} \\ &= \frac{0.1}{0.1 + 0.1} \\ &= 0.5 \end{aligned}$$

■

Example 2.4.4

The possibility of importing liquified natural gas (LNG) from Algeria has been suggested as one way of coping with a future energy crunch. Complicating matters, though, is the fact that LNG is highly volatile and poses an enormous safety hazard. Any major spill occurring near a U.S. port could result in a fire of catastrophic proportions. The question, therefore, of the *likelihood* of a spill becomes critical input for future policymakers who may have to decide whether or not to implement the proposal.

Two numbers need to be taken into account: (1) the probability that a tanker will have an accident near a port, and (2) the probability that a major spill will develop *given* that an accident has happened. Although no significant spills of LNG have yet occurred anywhere in the world, these probabilities can be approximated from records kept on similar tankers transporting less dangerous cargo. On the basis of such data, it has been estimated (47) that the probability is 8/50,000 that an LNG tanker will have an accident on any one trip. Given that an accident *has* occurred, it is suspected that only three times in fifteen thousand will the damage be sufficiently severe that a major spill would develop. What are the chances that a given LNG shipment would precipitate a catastrophic disaster?

Let A denote the event "Spill develops" and let B denote the event "Accident occurs." Past experience is suggesting that $P(B) = 8/50{,}000$ and $P(A|B) = 3/15{,}000$. Of primary concern is the probability that an accident will occur *and* a spill will ensue—that is, $P(A \cap B)$. Using Equation 2.4.2, we find that the chances of a catastrophic accident are on the order of three in one hundred million:

$$\begin{aligned} P(\text{Accident occurs and spill develops}) &= P(A \cap B) \\ &= P(A|B)P(B) \\ &= \frac{3}{15{,}000} \cdot \frac{8}{50{,}000} \\ &= 0.000000032 \end{aligned}$$

■

Example 2.4.5

Max and Muffy are two myopic deer hunters who shoot simultaneously at a nearby sheepdog that they have mistaken for a 10-point buck. Based on years of well-documented ineptitude, it can be assumed that Max has a 20% chance of hitting a stationary target at close range, Muffy has a 30% chance, and the probability is 0.06 that they will both be on target. Suppose that the sheepdog is hit and killed by exactly one bullet. What is the probability that Muffy fired the fatal shot?

Let A be the event that Max hit the dog, and let B be the event that Muffy hit the dog. Then $P(A) = 0.2$, $P(B) = 0.3$, and $P(A \cap B) = 0.06$. We are trying to find

$$P(B|(A^C \cap B) \cup (A \cap B^C))$$

where the event $(A^C \cap B) \cup (A \cap B^C)$ is the union of A and B minus the intersection—that is, it represents the event that either A or B but not both occur (recall Figure 2.4.4).

Notice, also, from Figure 2.4.4 that the intersection of B and $(A^C \cap B) \cup (A \cap B^C)$ is the event $A^C \cap B$. Therefore, from Definition 2.4.1,

$$\begin{aligned} P(B|(A^C \cap B) \cup (A \cap B^C)) &= [P(A^C \cap B)]/[P\{(A^C \cap B) \cup (A \cap B^C)\}] \\ &= [P(B) - P(A \cap B)]/[P(A \cup B) - P(A \cap B)] \\ &= [0.3 - 0.06]/[0.2 + 0.3 - 0.06 - 0.06] \\ &= 0.63 \end{aligned}$$

■

CASE STUDY 2.4.1

Several years ago, a television program (inadvertently) spawned a conditional probability problem that led to more than a few heated discussions, even in the national media. The show was *Let's Make a Deal*, and the question involved the strategy that contestants should take to maximize their chances of winning prizes.

On the program, a contestant would be presented with three doors, behind one of which was the prize. After the contestant had selected a door, the host, Monty Hall, would open one of the other two doors, showing that the prize was not there. Then he would give the contestant a choice—either stay with the door initially selected or switch to the "third" door, which had not been opened.

For many viewers, common sense seemed to suggest that switching doors would make no difference. By assumption, the prize had a one-third chance of being behind each of the doors when the game began. Once a door was opened, it was argued that each of the remaining doors now had a one-half probability of hiding the prize, so contestants gained nothing by switching their bets.

Not so. An application of Definition 2.4.1 shows that it *did* make a difference—contestants, in fact, *doubled* their chances of winning by switching doors. To see why, consider a specific (but typical) case: The contestant has bet on Door #2 and Monty Hall has opened Door #3. Given that sequence of events, we need to calculate and compare the conditional probability of the prize being behind Door #1 and Door #2, respectively. If the former is larger (and we will prove that it is), the contestant should switch doors.

Table 2.4.1 shows the sample space associated with the scenario just described. If the prize is actually behind Door #1, the host has no choice but to open Door #3; similarly, if the prize is behind Door #3, the host has no choice but to open Door #1. In the event that the prize is behind Door #2, though, the host would (theoretically) open Door #1 half the time and Door #3 half the time.

Table 2.4.1

(Prize Location, Door Opened)	Probability
(1, 3)	1/3
(2, 1)	1/6
(2, 3)	1/6
(3, 1)	1/3

Notice that the four outcomes in S are not equally likely. There is necessarily a one-third probability that the prize is behind each of the three doors. However, the two choices that the host has when the prize is behind Door #2 necessitate that the two outcomes (2, 1) and (2, 3) share the one-third probability that represents the chances of the prize being behind Door #2. Each, then, has the one-sixth probability listed in Table 2.4.1.

Let A be the event that the prize is behind Door #2, and let B be the event that the host opened Door #3. Then

$$\begin{aligned} P(A|B) = P(\text{Contestant wins by not switching}) &= [P(A \cap B)]/P(B) \\ &= \left[\tfrac{1}{6}\right]/\left[\tfrac{1}{3} + \tfrac{1}{6}\right] \\ &= \tfrac{1}{3} \end{aligned}$$

Now, let A^* be the event that the prize is behind Door #1, and let B (as before) be the event that the host opens Door #3. In this case,

$$\begin{aligned} P(A^*|B) = P(\text{Contestant wins by switching}) &= [P(A^* \cap B)]/P(B) \\ &= \left[\tfrac{1}{3}\right]/\left[\tfrac{1}{3} + \tfrac{1}{6}\right] \\ &= \tfrac{2}{3} \end{aligned}$$

Common sense would have led us astray again! If given the choice, contestants should have *always* switched doors. Doing so upped their chances of winning from one-third to two-thirds.

Questions

2.4.1. Suppose that two fair dice are tossed. What is the probability that the sum equals 10 given that it exceeds 8?

2.4.2. Find $P(A \cap B)$ if $P(A) = 0.2$, $P(B) = 0.4$, and $P(A|B) + P(B|A) = 0.75$.

2.4.3. If $P(A|B) < P(A)$, show that $P(B|A) < P(B)$.

2.4.4. Let A and B be two events such that $P((A \cup B)^C) = 0.6$ and $P(A \cap B) = 0.1$. Let E be the event that either A or B but not both will occur. Find $P(E|A \cup B)$.

2.4.5. Suppose that in Example 2.4.2 we ignored the ages of the children and distinguished only *three* family types: (boy, boy), (girl, boy), and (girl, girl). Would the conditional probability of both children being boys given that at least one is a boy be different from the answer found on pp. 33–34? Explain.

2.4.6. Two events, A and B, are defined on a sample space S such that $P(A|B) = 0.6$, P(At least one of the events occurs) $= 0.8$, and P(Exactly one of the events occurs) $= 0.6$. Find $P(A)$ and $P(B)$.

2.4.7. An urn contains one red chip and one white chip. One chip is drawn at random. If the chip selected is red, that chip together with two additional red chips are put back into the urn. If a white chip is drawn, the chip is returned to the urn. Then a second chip is drawn. What is the probability that both selections are red?

2.4.8. Given that $P(A) = a$ and $P(B) = b$, show that

$$P(A|B) \geq \frac{a + b - 1}{b}$$

2.4.9. An urn contains one white chip and a second chip that is equally likely to be white or black. A chip is drawn at random and returned to the urn. Then a second chip is drawn. What is the probability that a white appears on the second draw given that a white appeared on the first draw? (*Hint:* Let W_i be the event that a white chip is selected on the ith draw, $i = 1, 2$. Then $P(W_2|W_1) = \frac{P(W_1 \cap W_2)}{P(W_1)}$. If both chips in the urn are white, $P(W_1) = 1$; otherwise, $P(W_1) = \frac{1}{2}$.)

2.4.10. Suppose events A and B are such that $P(A \cap B) = 0.1$ and $P((A \cup B)^C) = 0.3$. If $P(A) = 0.2$, what does $P[(A \cap B)|(A \cup B)^C]$ equal? (*Hint:* Draw the Venn diagram.)

2.4.11. One hundred voters were asked their opinions of two candidates, A and B, running for mayor. Their responses to three questions are summarized below:

	Number Saying "Yes"
Do you like A?	65
Do you like B?	55
Do you like both?	25

(a) What is the probability that someone likes neither?

(b) What is the probability that someone likes exactly one?

(c) What is the probability that someone likes at least one?

(d) What is the probability that someone likes at most one?

(e) What is the probability that someone likes exactly one given that he or she likes at least one?

(f) Of those who like at least one, what proportion like both?

(g) Of those who do not like A, what proportion like B?

2.4.12. A fair coin is tossed three times. What is the probability that at least two heads will occur given that at most two heads have occurred?

2.4.13. Two fair dice are rolled. What is the probability that the number on the first die was at least as large as 4 given that the sum of the two dice was 8?

2.4.14. Four cards are dealt from a standard fifty-two-card poker deck. What is the probability that all four are aces given that at least three are aces? (*Note:* There are 270,725 different sets of four cards that can be dealt. Assume that the probability associated with each of those hands is 1/270,725.)

2.4.15. Given that $P(A \cap B^C) = 0.3$, $P((A \cup B)^C) = 0.2$, and $P(A \cap B) = 0.1$, find $P(A|B)$.

2.4.16. Given that $P(A) + P(B) = 0.9$, $P(A|B) = 0.5$, and $P(B|A) = 0.4$, find $P(A)$.

2.4.17. Let A and B be two events defined on a sample space S such that $P(A \cap B^C) = 0.1$, $P(A^C \cap B) = 0.3$, and $P((A \cup B)^C) = 0.2$. Find the probability that at least one of the two events occurs given that at most one occurs.

2.4.18. Suppose two dice are rolled. Assume that each possible outcome has probability 1/36. Let A be the event that the sum of the two dice is greater than or equal to 8, and let B be the event that at least one of the dice shows a 5. Find $P(A|B)$.

2.4.19. According to your neighborhood bookie, five horses are scheduled to run in the third race at the local track, and handicappers have assigned them the following probabilities of winning:

Horse	Probability of Winning
Scorpion	0.10
Starry Avenger	0.25
Australian Doll	0.15
Dusty Stake	0.30
Outandout	0.20

Suppose that Australian Doll and Dusty Stake are scratched from the race at the last minute. What are the chances that Outandout will prevail over the reduced field?

2.4.20. Andy, Bob, and Charley have all been serving time for grand theft auto. According to prison scuttlebutt, the warden plans to release two of the three next week. They all have identical records, so the two to be released will be chosen at random, meaning that each has a two-thirds probability of being included in the two to be set free. Andy, however, is friends with a guard who will know ahead of time which two will leave. He offers to tell Andy the name of one prisoner *other than himself* who will be released. Andy, however, declines the offer, believing that if he learns the name of one prisoner scheduled to be released, then *his* chances of being the other person set free will drop to one-half (since only two prisoners will be left at that point). Is his concern justified?

APPLYING CONDITIONAL PROBABILITY TO HIGHER-ORDER INTERSECTIONS

We have seen that conditional probabilities can be useful in evaluating intersection probabilities—that is, $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$. A similar result holds for higher-order intersections. Consider $P(A \cap B \cap C)$. By thinking of $A \cap B$ as a single event—say, D—we can write

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

Repeating this same argument for n events, $A_1, A_2, \ldots, A_n$, gives a formula for the general case:

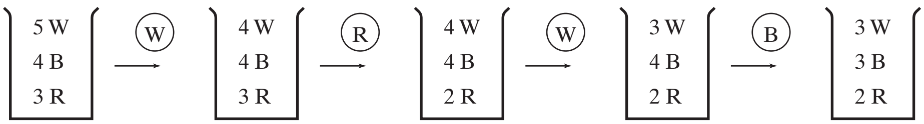
$$\begin{aligned} P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \\ \cdot P(A_{n-1}|A_1 \cap A_2 \cap \cdots \cap A_{n-2}) \cdots P(A_2|A_1) \cdot P(A_1) \end{aligned} \tag{2.4.3}$$

Example 2.4.6

An urn contains five white chips, four black chips, and three red chips. Four chips are drawn sequentially and without replacement. What is the probability of obtaining the sequence (white, red, white, black)?

Figure 2.4.5. shows the evolution of the urn's composition as the desired sequence is assembled. Define the following four events:

Figure 2.4.5



A: white chip is drawn on first selection
B: red chip is drawn on second selection
C: white chip is drawn on third selection
D: black chip is drawn on fourth selection

Our objective is to find $P(A \cap B \cap C \cap D)$.

From Equation 2.4.3,

$$P(A \cap B \cap C \cap D) = P(D|A \cap B \cap C) \cdot P(C|A \cap B) \cdot P(B|A) \cdot P(A)$$

Each of the probabilities on the right-hand side of the equation here can be gotten by just looking at the urns pictured in Figure 2.4.5: $P(D|A \cap B \cap C) = \frac{4}{9}$, $P(C|A \cap B) = \frac{4}{10}$, $P(B|A) = \frac{3}{11}$, and $P(A) = \frac{5}{12}$. Therefore, the probability of drawing a (white, red, white, black) sequence is 0.02:

$$\begin{aligned} P(A \cap B \cap C \cap D) &= \frac{4}{9} \cdot \frac{4}{10} \cdot \frac{3}{11} \cdot \frac{5}{12} \\ &= \frac{240}{11{,}880} \\ &= 0.02 \end{aligned}$$

■

CASE STUDY 2.4.2

Since the late 1940s, tens of thousands of eyewitness accounts of strange lights in the skies, unidentified flying objects, and even alleged abductions by little green men have made headlines. None of these incidents, though, has produced any hard evidence, any irrefutable *proof* that Earth has been visited by a race of extraterrestrials. Still, the haunting question remains—are we alone in the universe? Or are there other civilizations, more advanced than ours, making the occasional flyby?

Until, or unless, a flying saucer plops down on the White House lawn and a strange-looking creature emerges with the proverbial "Take me to your leader" demand, we may never know whether we have any cosmic neighbors. Equation 2.4.3, though, can help us speculate on the *probability* of our not being alone.

Recent discoveries suggest that planetary systems much like our own may be quite common. If so, there are likely to be many planets whose chemical makeups, temperatures, pressures, and so on, are suitable for life. Let those planets be the points in our sample space. Relative to them, we can define three events:

A: life arises
B: technical civilization arises (one capable of interstellar communication)
C: technical civilization is flourishing *now*

(Continued on next page)

(Case Study 2.4.2 continued)

In terms of A, B, and C, the probability that a habitable planet is presently supporting a technical civilization is the probability of an intersection—specifically, $P(A \cap B \cap C)$. Associating a number with $P(A \cap B \cap C)$ is highly problematic, but the task is simplified considerably if we work instead with the equivalent conditional formula, $P(C|B \cap A) \cdot P(B|A) \cdot P(A)$.

Scientists speculate (163) that life of some kind may arise on one-third of all planets having a suitable environment and that life on maybe 1% of all those planets will evolve into a technical civilization. In our notation, $P(A) = \frac{1}{3}$ and $P(B|A) = \frac{1}{100}$.

More difficult to estimate is $P(C|A \cap B)$. On Earth, we have had the capability of interstellar communication (that is, radio astronomy) for only a few decades, so $P(C|A \cap B)$, *empirically*, is on the order of 1×10^{-8}. But that may be an overly pessimistic estimate of a technical civilization's ability to endure. It may be true that if a civilization can avoid annihilating itself when it first develops nuclear weapons, its prospects for longevity are fairly good. If that were the case, $P(C|A \cap B)$ might be as large as 1×10^{-2}.

Putting these estimates into the computing formula for $P(A \cap B \cap C)$ yields a range for the probability of a habitable planet currently supporting a technical civilization. The chances may be as small as 3.3×10^{-11} or as "large" as 3.3×10^{-5}:

$$(1 \times 10^{-8})\left(\frac{1}{100}\right)\left(\frac{1}{3}\right) < P(A \cap B \cap C) < (1 \times 10^{-2})\left(\frac{1}{100}\right)\left(\frac{1}{3}\right)$$

or

$$0.000000000033 < P(A \cap B \cap C) < 0.000033$$

A better way to put these figures in some kind of perspective is to think in terms of *numbers* rather than probabilities. Astronomers estimate there are 3×10^{11} habitable planets in our Milky Way galaxy. Multiplying that total by the two limits for $P(A \cap B \cap C)$ gives an indication of *how many* cosmic neighbors we are likely to have. Specifically, $3 \times 10^{11} \cdot 0.000000000033 \doteq 10$, while $3 \times 10^{11} \cdot 0.000033 \doteq 10{,}000{,}000$. So, on the one hand, we may be a galactic rarity. At the same time, the probabilities do not preclude the very real possibility that the Milky Way is abuzz with activity and that our neighbors number in the millions.

About the Data In 2009, NASA launched the spacecraft Kepler, whose sole mission was to find other solar systems in the Milky Way galaxy. And find them it did! By 2015 it had documented over one thousand "exoplanets," including more than a dozen belonging to so-called *Goldilocks zones*, meaning the planet's size and configuration of its orbit would allow for liquid water to remain on its surface, thus making it a likely home for living organisms. Of particular interest is the planet Kepler 186f located in the constellation Cygnus, some five hundred light years away. Its size, orbit, and nature of its sun make it remarkably similar to Earth—maybe not quite a *twin*, astronomers say, but certainly a *cousin*. (So, maybe UFOs are real and maybe a few have landed, and maybe that weird guy your sister is dating isn't really from Cleveland. . .)

Questions

2.4.21. An urn contains six white chips, four black chips, and five red chips. Five chips are drawn out, one at a time and without replacement. What is the probability of getting the sequence (black, black, red, white, white)? Suppose that the chips are numbered 1 through 15. What is the probability of getting a specific sequence—say, (2, 6, 4, 9, 13)?

2.4.22. A man has n keys on a key ring, one of which opens the door to his apartment. Having celebrated a bit too much one evening, he returns home only to find himself unable to distinguish one key from another. Resourceful, he works out a fiendishly clever plan: He will choose a key at random and try it. If it fails to open the door, he will discard it and choose at random one of the remaining $n - 1$ keys, and so on. Clearly, the probability that he gains entrance with the first key he selects is $1/n$. Show that the probability the door opens with the *third* key he tries is also $1/n$. (*Hint:* What has to happen before he even gets to the third key?)

2.4.23. Your favorite college football team has had a good season so far but they need to win at least two of their last four games to qualify for a New Year's Day bowl bid. Oddsmakers estimate the team's probabilities of winning each of their last four games to be 0.60, 0.50, 0.40, and 0.70, respectively.

(a) What are the chances that you will get to watch your team play on Jan. 1?

(b) Is the probability that your team wins all four games given that they have won at least three games equal to the probability that they win the fourth game? Explain.

(c) Is the probability that your team wins all four games given that they won the first three games equal to the probability that they win the fourth game?

2.4.24. One chip is drawn at random from an urn that contains one white chip and one black chip. If the white chip is selected, we simply return it to the urn; if the black chip is drawn, that chip—together with another black—are returned to the urn. Then a second chip is drawn, with the same rules for returning it to the urn. Calculate the probability of drawing two whites followed by three blacks.

CALCULATING "UNCONDITIONAL" AND "INVERSE" PROBABILITIES

We conclude this section with two very useful theorems that apply to *partitioned* sample spaces. By definition, a set of events $A_1, A_2, \ldots, A_n$ "partition" S if every outcome in the sample space belongs to one and only one of the A_i's—that is, the A_i's are mutually exclusive and their union is S (see Figure 2.4.6).

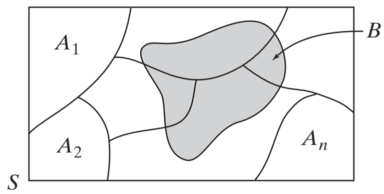


Figure 2.4.6

Let B, as pictured, denote any event defined on S. The first result, Theorem 2.4.1, gives a formula for the "unconditional" probability of B (in terms of the A_i's). Then Theorem 2.4.2 calculates the set of conditional probabilities, $P(A_j|B)$, $j = 1, 2, \ldots, n$.

Theorem 2.4.1 *Let $\{A_i\}_{i=1}^n$ be a set of events defined over S such that $S = \bigcup_{i=1}^n A_i$, $A_i \cap A_j = \emptyset$ for $i \neq j$, and $P(A_i) > 0$ for $i = 1, 2, \ldots, n$. For any event B,*

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

Proof By the conditions imposed on the A_i's,

$$B = (B \cap A_1) \cup (B \cap A_2) \cup \cdots \cup (B \cap A_n)$$

and

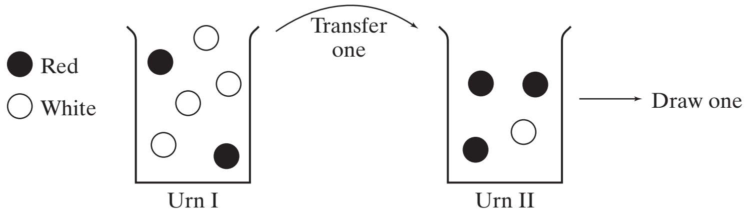
$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \cdots + P(B \cap A_n)$$

But each $P(B \cap A_i)$ can be written as the product $P(B|A_i)P(A_i)$, and the result follows.

Example 2.4.7 Urn I contains two red chips and four white chips; urn II, three red and one white. A chip is drawn at random from urn I and transferred to urn II. Then a chip is drawn from urn II. What is the probability that the chip drawn from urn II is red?

Let B be the event "Chip drawn from urn II is red"; let A_1 and A_2 be the events "Chip transferred from urn I is red" and "Chip transferred from urn I is white," respectively. By inspection (see Figure 2.4.7), we can deduce all the probabilities appearing in the right-hand side of the formula in Theorem 2.4.1:

Figure 2.4.7



$$P(B|A_1) = \frac{4}{5} \quad P(B|A_2) = \frac{3}{5}$$
$$P(A_1) = \frac{2}{6} \quad P(A_2) = \frac{4}{6}$$

Putting all this information together, we see that the chances are two out of three that a red chip will be drawn from urn II:

$$\begin{aligned} P(B) &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) \\ &= \frac{4}{5} \cdot \frac{2}{6} + \frac{3}{5} \cdot \frac{4}{6} \\ &= \frac{2}{3} \end{aligned}$$

■

Example 2.4.8 A standard poker deck is shuffled, and the card on top is removed. What is the probability that the *second* card is an ace?

Define the following events:

B: second card is an ace
A_1: top card was an ace
A_2: top card was not an ace

Then $P(B|A_1) = \frac{3}{51}$, $P(B|A_2) = \frac{4}{51}$, $P(A_1) = \frac{4}{52}$, and $P(A_2) = \frac{48}{52}$. Since the A_i's partition the sample space of two-card selections, Theorem 2.4.1 applies. Substituting into the expression for $P(B)$ shows that $\frac{4}{52}$ is the probability that the second card is an ace:

$$\begin{aligned} P(B) &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) \\ &= \frac{3}{51} \cdot \frac{4}{52} + \frac{4}{51} \cdot \frac{48}{52} \\ &= \frac{4}{52} \end{aligned}$$

Comment Notice that $P(B) = P(\text{second card is an ace})$ is numerically the same as $P(A_1) = P(\text{first card is an ace})$. The analysis in Example 2.4.8 illustrates a basic principle in probability that says, in effect, "What you don't know, doesn't matter." Here, removal of the top card is irrelevant to any subsequent probability calculations *if the identity of that card remains unknown.* ■

Example 2.4.9 Ashley is hoping to land a summer internship with a public relations firm. If her interview goes well, she has a 70% chance of getting an offer. If the interview is a bust, though, her chances of getting the position drop to 20%. Unfortunately, Ashley tends to babble incoherently when she is under stress, so the likelihood of the interview going well is only 0.10. What is the probability that Ashley gets the internship?

Let B be the event "Ashley is offered internship," let A_1 be the event "Interview goes well," and let A_2 be the event "Interview does not go well." By assumption,

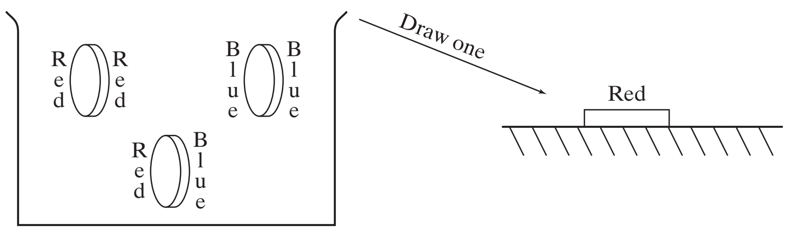
$$\begin{array}{ll} P(B|A_1) = 0.70 & P(B|A_2) = 0.20 \\ P(A_1) \quad = 0.10 & P(A_2) \quad = 1 - P(A_1) = 1 - 0.10 = 0.90 \end{array}$$

According to Theorem 2.4.1, Ashley has a *25%* chance of landing the internship:

$$\begin{aligned} P(B) &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) \\ &= (0.70)(0.10) + (0.20)(0.90) \\ &= 0.25 \end{aligned}$$ ■

Example 2.4.10 Three chips are placed in an urn. One is red on both sides, a second is blue on both sides, and the third is red on one side and blue on the other. One chip is selected at random and placed on a table. Suppose that the color showing on that chip is red. What is the probability that the color underneath is also red (see Figure 2.4.8)?

Figure 2.4.8



At first glance, it may seem that the answer is one-half: We know that the blue/blue chip has not been drawn, and only one of the remaining two—the red/red chip—satisfies the event that the color underneath is red. If this game were played over and over, though, and records were kept of the outcomes, it would be found that the proportion of times that a red top has a red bottom is two-thirds, not the one-half that our intuition might suggest. The correct answer follows from an application of Theorem 2.4.1.

Define the following events:

A: bottom side of chip drawn is red
B: top side of chip drawn is red
A_1: red/red chip is drawn
A_2: blue/blue chip is drawn
A_3: red/blue chip is drawn

From the definition of conditional probability,

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

But $P(A \cap B) = P(\text{Both sides are red}) = P$ (red/red chip) $= \frac{1}{3}$, and Theorem 2.4.1 can be used to find the denominator, $P(B)$:

$$\begin{aligned} P(B) &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3) \\ &= 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} \\ &= \frac{1}{2} \end{aligned}$$

Therefore,

$$P(A|B) = \frac{1/3}{1/2} = \frac{2}{3}$$

■

Comment The question posed in Example 2.4.10 gives rise to a simple but effective con game. The trick is to convince a "mark" that the initial analysis given earlier is correct, meaning that the bottom has a fifty-fifty chance of being the same color as the top. Under that incorrect presumption that the game is "fair," both participants put up the same amount of money, but the gambler (knowing the correct analysis) always bets that the bottom is the same color as the top. In the long run, then, the con artist will be winning an even-money bet two-thirds of the time!

Questions

2.4.25. A toy manufacturer buys ball bearings from three different suppliers—50% of her total order comes from supplier 1, 30% from supplier 2, and the rest from supplier 3. Past experience has shown that the quality-control standards of the three suppliers are not all the same. Two percent of the ball bearings produced by supplier 1 are defective, while suppliers 2 and 3 produce defective bearings 3% and 4% of the time, respectively. What proportion of the ball bearings in the toy manufacturer's inventory are defective?

2.4.26. A fair coin is tossed. If a head turns up, a fair die is tossed; if a tail turns up, two fair dice are tossed. What is the probability that the face (or the sum of the faces) showing on the die (or the dice) is equal to 6?

2.4.27. Foreign policy experts estimate that the probability is 0.65 that war will break out next year between two Middle East countries if either side significantly escalates its terrorist activities. Otherwise, the likelihood of war is estimated to be 0.05. Based on what has happened this year, the chances of terrorism reaching a critical level in the next twelve months are thought to be three in ten. What is the probability that the two countries will go to war?

2.4.28. A telephone solicitor is responsible for canvassing three suburbs. In the past, 60% of the completed calls to Belle Meade have resulted in contributions, compared to 55% for Oak Hill and 35% for Antioch. Her list of telephone numbers includes one thousand households from Belle Meade, one thousand from Oak Hill, and two thousand from Antioch. Suppose that she picks a number at random from the list and places the call. What is the probability that she gets a donation?

2.4.29. If men constitute 47% of the population and tell the truth 78% of the time, while women tell the truth 63% of the time, what is the probability that a person selected at random will answer a question truthfully?

2.4.30. Urn I contains three red chips and one white chip. Urn II contains two red chips and two white chips. One chip is drawn from each urn and transferred to the other urn. Then a chip is drawn from the first urn. What is the probability that the chip ultimately drawn from urn I is red?

2.4.31. Medical records show that 0.01% of the general adult population not belonging to a high-risk group (for example, intravenous drug users) are HIV-positive. Blood tests for the virus are 99.9% accurate when given to someone infected and 99.99% accurate when given to someone not infected. What is the probability that a random adult not in a high-risk group will test positive for the HIV virus?

2.4.32. Recall the "survival" lottery described in Question 2.2.14. What is the probability of release associated with the prisoner's optimal strategy?

2.4.33. In an upstate congressional race, the incumbent Republican (R) is running against a field of three Democrats (D_1, D_2, and D_3) seeking the nomination. Political pundits estimate that the probabilities of D_1, D_2, or D_3 winning the primary are 0.35, 0.40, and 0.25, respectively. Furthermore, results from a variety of polls are suggesting that R would have a 40% chance of defeating D_1 in the general election, a 35% chance of defeating D_2, and a 60% chance of defeating D_3. Assuming all these estimates to be accurate, what are the chances that the Republican will retain his seat?

2.4.34. An urn contains forty red chips and sixty white chips. Six chips are drawn out and discarded, and a seventh chip is drawn. What is the probability that the seventh chip is red?

2.4.35. A study has shown that seven out of ten people will say "heads" if asked to call a coin toss. Given that the coin is fair, though, a head occurs, on the average, only five times out of ten. Does it follow that you have the advantage if you let the other person call the toss? Explain.

2.4.36. Based on pretrial speculation, the probability that a jury returns a guilty verdict in a certain high-profile murder case is thought to be 15% if the defense can discredit the police department and 80% if they cannot. Veteran court observers believe that the skilled defense attorneys have a 70% chance of convincing the jury that the police either contaminated or planted some of the key evidence. What is the probability that the jury returns a guilty verdict?

2.4.37. As an incoming freshman, Marcus believes that he has a 25% chance of earning a GPA in the 3.5 to 4.0 range, a 35% chance of graduating with a 3.0 to 3.5 GPA, and a 40% chance of finishing with a GPA less than 3.0. From what the pre-med advisor has told him, Marcus has an eight in ten chance of getting into medical school if his GPA is above 3.5, a five in ten chance if his GPA is in the 3.0 to 3.5 range, and only a one in ten chance if his GPA falls below 3.0. Based on those estimates, what is the probability that Marcus gets into medical school?

2.4.38. The governor of a certain state has decided to come out strongly for prison reform and is preparing a new early release program. Its guidelines are simple: prisoners related to members of the governor's staff would have a 90% chance of being released early; the probability of early release for inmates not related to the governor's staff would be 0.01. Suppose that 40% of all inmates are related to someone on the governor's staff. What is the probability that a prisoner selected at random would be eligible for early release?

2.4.39. Following are the percentages of students of State College enrolled in each of the school's main divisions. Also listed are the proportions of students in each division who are women.

Division	%	% Women
Humanities	40	60
Natural science	10	15
History	30	45
Social science	20	75
	100	

Suppose the registrar selects one person at random. What is the probability that the student selected will be a male?

BAYES' THEOREM

The second result in this section that is set against the backdrop of a partitioned sample space has a curious history. The first explicit statement of Theorem 2.4.2, coming in 1812, was due to Laplace, but it was named after the Reverend Thomas Bayes, whose 1763 paper (published posthumously) had already outlined the result. On one level, the theorem is a relatively minor extension of the definition of conditional probability. When viewed from a loftier perspective, though, it takes on some rather profound philosophical implications. The latter, in fact, have precipitated a schism among practicing statisticians: "Bayesians" analyze data one way; "non-Bayesians" often take a fundamentally different approach (see Section 5.8).

Our use of the result here will have nothing to do with its statistical interpretation. We will apply it simply as the Reverend Bayes originally intended, as a formula for evaluating a certain kind of "inverse" probability. If we know $P(B|A_i)$ for all i, the theorem enables us to compute conditional probabilities "in the other direction"—that is, we can deduce $P(A_j|B)$ from the $P(B|A_i)$'s.

Theorem 2.4.2 *(Bayes') Let $\{A_i\}_{i=1}^n$ be a set of n events, each with positive probability, that partition S in such a way that $\cup_{i=1}^n A_i = S$ and $A_i \cap A_j = \emptyset$ for $i \neq j$. For any event B (also defined on S), where $P(B) > 0$,*

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$$

for any $1 \leq j \leq n$.

Proof From Definition 2.4.1,

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j)P(A_j)}{P(B)}$$

But Theorem 2.4.1 allows the denominator to be written as $\sum_{i=1}^{n} P(B|A_i)P(A_i)$, and the result follows.

Problem-Solving Hints

(Working with Partitioned Sample Spaces)

Students sometimes have difficulty setting up problems that involve partitioned sample spaces—in particular, ones whose solution requires an application of either Theorem 2.4.1 or 2.4.2—because of the nature and amount of information that need to be incorporated into the answers. The "trick" is learning to identify which part of the "given" corresponds to B and which parts correspond to the A_i's. The following hints may help.

1. As you read the question, pay particular attention to the last one or two sentences. Is the problem asking for an *unconditional probability* (in which case Theorem 2.4.1 applies) or a *conditional probability* (in which case Theorem 2.4.2 applies)?
2. If the question is asking for an unconditional probability, let B denote the event whose probability you are trying to find; if the question is asking for a conditional probability, let B denote the event that has *already happened*.
3. Once event B has been identified, reread the beginning of the question and assign the A_i's.

Example 2.4.11 A biased coin, twice as likely to come up heads as tails, is tossed once. If it shows heads, a chip is drawn from urn I, which contains three white chips and four red chips; if it shows tails, a chip is drawn from urn II, which contains six white chips and three red chips. Given that a white chip was drawn, what is the probability that the coin came up tails (see Figure 2.4.9)?

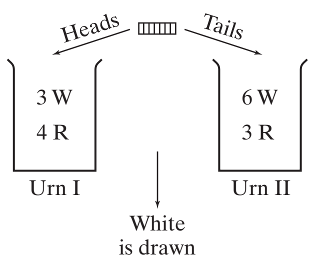


Figure 2.4.9

Since $P(\text{heads}) = 2P(\text{tails})$, it must be true that $P(\text{heads}) = \frac{2}{3}$ and $P(\text{tails}) = \frac{1}{3}$. Define the events

B: white chip is drawn
A_1: coin came up heads (i.e., chip came from urn I)
A_2: coin came up tails (i.e., chip came from urn II)

Our objective is to find $P(A_2|B)$. From Figure 2.4.9,

$$P(B|A_1) = \frac{3}{7} \quad P(B|A_2) = \frac{6}{9}$$
$$P(A_1) = \frac{2}{3} \quad P(A_2) = \frac{1}{3}$$

so

$$\begin{aligned} P(A_2|B) &= \frac{P(B|A_2)P(A_2)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)} \\ &= \frac{(6/9)(1/3)}{(3/7)(2/3) + (6/9)(1/3)} \\ &= \frac{7}{16} \end{aligned}$$

■

Example 2.4.12

During a power blackout, one hundred persons are arrested on suspicion of looting. Each is given a polygraph test. From past experience it is known that the polygraph is 90% reliable when administered to a guilty suspect and 98% reliable when given to someone who is innocent. Suppose that of the one hundred persons taken into custody, only twelve were actually involved in any wrongdoing. What is the probability that a given suspect is innocent given that the polygraph says he is guilty?

Let B be the event "Polygraph says suspect is guilty," and let A_1 and A_2 be the events "Suspect is guilty" and "Suspect is not guilty," respectively. To say that the polygraph is "90% reliable when administered to a guilty suspect" means that $P(B|A_1) = 0.90$. Similarly, the 98% reliability for innocent suspects implies that $P(B^C|A_2) = 0.98$, or, equivalently, $P(B|A_2) = 0.02$.

We also know that $P(A_1) = \frac{12}{100}$ and $P(A_2) = \frac{88}{100}$. Substituting into Theorem 2.4.2, then, shows that the probability a suspect is innocent given that the polygraph says he is guilty is 0.14:

$$\begin{aligned} P(A_2|B) &= \frac{P(B|A_2)P(A_2)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)} \\ &= \frac{(0.02)(88/100)}{(0.90)(12/100) + (0.02)(88/100)} \\ &= 0.14 \end{aligned}$$

■

Example 2.4.13

As medical technology advances and adults become more health conscious, the demand for diagnostic screening tests inevitably increases. Looking for problems, though, when no symptoms are present can have undesirable consequences that may outweigh the intended benefits.

Suppose, for example, a woman has a medical procedure performed to see whether she has a certain type of cancer. Let B denote the event that the test *says* she has cancer, and let A_1 denote the event that she actually *does* (and A_2, the event that she *does not*). Furthermore, suppose the prevalence of the disease and the precision of the diagnostic test are such that

$$P(A_1) = 0.0001 \qquad [\text{and} \quad P(A_2) = 0.9999]$$

$$P(B|A_1) = 0.90 = P(\text{Test } \textit{says} \text{ woman has cancer when, in fact, she does})$$

$$P(B|A_2) = P\left(B|A_1^C\right) = 0.001 = P(\text{False positive}) = P(\text{Test } \textit{says} \text{ woman has cancer when, in fact, she does not})$$

What is the probability that she *does* have cancer, given that the diagnostic procedure says she does? That is, calculate $P(A_1|B)$.

Although the method of solution here is straightforward, the actual numerical answer is not what we would expect. From Theorem 2.4.2,

$$\begin{aligned} P(A_1|B) &= \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_1^C)P(A_1^C)} \\ &= \frac{(0.9)(0.0001)}{(0.9)(0.0001) + (0.001)(0.9999)} \\ &= 0.08 \end{aligned}$$

So, only 8% of those women identified as having cancer actually do! Table 2.4.2 shows the strong dependence of $P(A_1|B)$ on $P(A_1)$ and $P(B|A_1^C)$.

Table 2.4.2

$P(A_1)$	$P(B\|A_1^C)$	$P(A_1\|B)$
0.0001	0.001	0.08
	0.0001	0.47
0.001	0.001	0.47
	0.0001	0.90
0.01	0.001	0.90
	0.0001	0.99

In light of these probabilities, the practicality of screening programs directed at diseases having a low prevalence is open to question, especially when the diagnostic procedure, itself, poses a nontrivial health risk. (For precisely those two reasons, the use of chest X-rays to screen for tuberculosis is no longer advocated by the medical community.) ■

Example 2.4.14

According to the manufacturer's specifications, your home burglar alarm has a 95% chance of going off if someone breaks into your house. During the two years you have lived there, the alarm has gone off on five different nights, each time for no apparent reason. Suppose the alarm goes off tomorrow night. What is the probability that someone is trying to break into your house? (*Note:* Police statistics show that the chances of any particular house in your neighborhood being burglarized on any given night are two in ten thousand.)

Let B be the event "Alarm goes off tomorrow night," and let A_1 and A_2 be the events "House is being burglarized" and "House is not being burglarized," respectively. Then

$$\begin{aligned} P(B|A_1) &= 0.95 \\ P(B|A_2) &= 5/730 \quad \text{(i.e., five nights in two years)} \\ P(A_1) &= 2/10,000 \\ P(A_2) &= 1 - P(A_1) = 9998/10,000 \end{aligned}$$

The probability in question is $P(A_1|B)$.

Intuitively, it might seem that $P(A_1|B)$ should be close to 1 because the alarm's "performance" probabilities look good—$P(B|A_1)$ is close to 1 (as it should be) and

$P(B|A_2)$ is close to 0 (as it should be). Nevertheless, $P(A_1|B)$ turns out to be surprisingly small:

$$\begin{aligned}P(A_1|B) &= \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)}\\ &= \frac{(0.95)(2/10{,}000)}{(0.95)(2/10{,}000) + (5/730)(9998/10{,}000)}\\ &= 0.027\end{aligned}$$

That is, if you hear the alarm going off, the probability is only 0.027 that your house is being burglarized.

Computationally, the reason $P(A_1|B)$ is so small is that $P(A_2)$ is so large. The latter makes the denominator of $P(A_1|B)$ large and, in effect, "washes out" the numerator. Even if $P(B|A_1)$ were substantially increased (by installing a more expensive alarm), $P(A_1|B)$ would remain largely unchanged (see Table 2.4.3).

Table 2.4.3

	$P(B\|A_1)$			
	0.95	0.97	0.99	0.999
$P(A_1\|B)$	0.027	0.028	0.028	0.028

■

Questions

2.4.40. Urn I contains two white chips and one red chip; urn II has one white chip and two red chips. One chip is drawn at random from urn I and transferred to urn II. Then one chip is drawn from urn II. Suppose that a red chip is selected from urn II. What is the probability that the chip transferred was white?

2.4.41. Urn I contains three red chips and five white chips; urn II contains four reds and four whites; urn III contains five reds and three whites. One urn is chosen at random and one chip is drawn from that urn. Given that the chip drawn was red, what is the probability that III was the urn sampled?

2.4.42. A dashboard warning light is supposed to flash red if a car's oil pressure is too low. On a certain model, the probability of the light flashing when it should is 0.99; 2% of the time, though, it flashes for no apparent reason. If there is a 10% chance that the oil pressure really is low, what is the probability that a driver needs to be concerned if the warning light goes on?

2.4.43. Building permits were issued last year to three contractors starting up a new subdivision: Tara Construction built two houses; Westview, three houses; and Hearthstone, six houses. Tara's houses have a 60% probability of developing leaky basements; homes built by Westview and Hearthstone have that same problem 50% of the time and 40% of the time, respectively. Yesterday, the Better Business Bureau received a complaint from one of the new homeowners that his basement is leaking. Who is most likely to have been the contractor?

2.4.44. Two sections of a senior probability course are being taught. From what she has heard about the two instructors listed, Francesca estimates that her chances of passing the course are 0.85 if she gets Professor X and 0.60 if she gets Professor Y. The section into which she is put is determined by the registrar. Suppose that her chances of being assigned to Professor X are four out of ten. Fifteen weeks later we learn that Francesca did, indeed, pass the course. What is the probability she was enrolled in Professor X's section?

2.4.45. A liquor store owner is willing to cash personal checks for amounts up to $50, but she has become wary of customers who wear sunglasses. Fifty percent of checks written by persons wearing sunglasses bounce. In contrast, 98% of the checks written by persons not wearing sunglasses clear the bank. She estimates that 10% of her customers wear sunglasses. If the bank returns a check and marks it "insufficient funds," what is the probability it was written by someone wearing sunglasses?

2.4.46. Brett and Margo have each thought about murdering their rich Uncle Basil in hopes of claiming their inheritance a bit early. Hoping to take advantage of Basil's predilection for immoderate desserts, Brett has put rat poison into the cherries flambé; Margo, unaware of Brett's activities, has laced the chocolate mousse with cyanide. Given the amounts likely to be eaten, the probability of the rat poison being fatal is 0.60; the cyanide, 0.90. Based on other dinners where Basil was presented with the same dessert options, we can assume that he has a 50% chance of asking for the cherries flambé, a 40% chance of ordering the chocolate mousse, and a 10% chance of skipping dessert altogether. No sooner are the dishes cleared away than Basil drops dead. In the absence of any other evidence, who should be considered the prime suspect?

2.4.47. Josh takes a twenty-question multiple-choice exam where each question has five possible answers. Some of the answers he knows, while others he gets right just by making lucky guesses. Suppose that the conditional probability of his knowing the answer to a randomly selected question given that he got it right is 0.92. How many of the twenty questions was he prepared for?

2.4.48. Recently the U.S. Senate Committee on Labor and Public Welfare investigated the feasibility of setting up a national screening program to detect child abuse. A team of consultants estimated the following probabilities: (1) one child in ninety is abused, (2) a screening program can detect an abused child 90% of the time, and (3) a screening program would incorrectly label 3% of all nonabused children as abused. What is the probability that a child is actually abused given that the screening program makes that diagnosis? How does the probability change if the incidence of abuse is one in one thousand? Or one in fifty?

2.4.49. At State University, 30% of the students are majoring in humanities, 50% in history and culture, and 20% in science. Moreover, according to figures released by the registrar, the percentages of women majoring in humanities, history and culture, and science are 75%, 45%, and 30%, respectively. Suppose Justin meets Anna at a fraternity party. What is the probability that Anna is a history and culture major?

2.4.50. An "eyes-only" diplomatic message is to be transmitted as a binary code of 0's and 1's. Past experience with the equipment being used suggests that if a 0 is sent, it will be (correctly) received as a 0 90% of the time (and mistakenly decoded as a 1 10% of the time). If a 1 is sent, it will be received as a 1 95% of the time (and as a 0 5% of the time). The text being sent is thought to be 70% 1's and 30% 0's. Suppose the next signal sent is received as a 1. What is the probability that it was sent as a 0?

2.4.51. When Zach wants to contact his girlfriend and he knows she is not at home, he is twice as likely to send her an e-mail as he is to leave a message on her phone. The probability that she responds to his e-mail within three hours is 80%; her chances of being similarly prompt in answering a phone message increase to 90%. Suppose she responded within two hours to the message he left this morning. What is the probability that Zach was communicating with her via e-mail?

2.4.52. A dot-com company ships products from three different warehouses (A, B, and C). Based on customer complaints, it appears that 3% of the shipments coming from A are somehow faulty, as are 5% of the shipments coming from B, and 2% coming from C. Suppose a customer is mailed an order and calls in a complaint the next day. What is the probability the item came from Warehouse C? Assume that Warehouses A, B, and C ship 30%, 20%, and 50% of the dot-com's sales, respectively.

2.4.53. A desk has three drawers. The first contains two gold coins, the second has two silver coins, and the third has one gold coin and one silver coin. A coin is drawn from a drawer selected at random. Suppose the coin selected was silver. What is the probability that the other coin in that drawer is gold?

2.4.54. An automobile insurance company has compiled the information summarized below on its policy-holders. Suppose someone calls to file a claim. To which age group does he or she most likely belong?

Age Group	% of Policyholders	% Involved in Accidents Last Year
Young (< 30)	20	35
Middle-aged (30–64)	50	15
Elderly (65+)	30	25

2.5 Independence

Section 2.4 dealt with the problem of reevaluating the probability of a given event in light of the additional information that some other event has already occurred. It often is the case, though, that the probability of the given event remains unchanged, regardless of the outcome of the second event—that is, $P(A|B) = P(A) = P(A|B^C)$. Events sharing this property are said to be *independent*. Definition 2.5.1 gives a necessary and sufficient condition for two events to be independent.

Definition 2.5.1
Two events A and B are said to be *independent* if $P(A \cap B) = P(A) \cdot P(B)$.

Comment The fact that the probability of the intersection of two independent events is equal to the product of their individual probabilities follows immediately from our first definition of independence, that $P(A|B) = P(A)$. Recall that the definition of conditional probability holds true for *any* two events A and B [provided that $P(B) > 0$]:

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

But $P(A|B)$ can equal $P(A)$ *only* if $P(A \cap B)$ factors into $P(A)$ times $P(B)$.

Example 2.5.1 Let A be the event of drawing a king from a standard poker deck and B, the event of drawing a diamond. Then, by Definition 2.5.1, A and B are independent because the probability of their intersection—drawing a king of diamonds—is equal to $P(A) \cdot P(B)$:

$$P(A \cap B) = \frac{1}{52} = \frac{1}{4} \cdot \frac{1}{13} = P(A) \cdot P(B)$$

■

Example 2.5.2 Suppose that A and B are independent events. Does it follow that A^C and B^C are also independent? That is, does $P(A \cap B) = P(A) \cdot P(B)$ guarantee that $P(A^C \cap B^C) = P(A^C) \cdot P(B^C)$?

Yes. The proof is accomplished by equating two different expressions for $P(A^C \cup B^C)$. First, by Theorem 2.3.6,

$$P(A^C \cup B^C) = P(A^C) + P(B^C) - P(A^C \cap B^C) \tag{2.5.1}$$

But the union of two complements is the complement of their intersection (recall Question 2.2.32). Therefore,

$$P(A^C \cup B^C) = 1 - P(A \cap B) \tag{2.5.2}$$

Combining Equations 2.5.1 and 2.5.2, we get

$$1 - P(A \cap B) = 1 - P(A) + 1 - P(B) - P(A^C \cap B^C)$$

Since A and B are independent, $P(A \cap B) = P(A) \cdot P(B)$, so

$$\begin{aligned} P(A^C \cap B^C) &= 1 - P(A) + 1 - P(B) - [1 - P(A) \cdot P(B)] \\ &= [1 - P(A)][1 - P(B)] \\ &= P(A^C) \cdot P(B^C) \end{aligned}$$

the latter factorization implying that A^C and B^C are, themselves, independent. ■

Example 2.5.3 Electronics Warehouse is responding to affirmative-action litigation by establishing hiring goals by race and sex for its office staff. So far they have agreed to employ the 120 people characterized in Table 2.5.1. How many black women do they need in order for the events A: Employee is female and B: Employee is black to be independent?

Let x denote the number of black women necessary for A and B to be independent. Then

$$P(A \cap B) = P(\text{Black female}) = x/(120 + x)$$

must equal

$$P(A)P(B) = P(\text{female})P(\text{black}) = [(40 + x)/(120 + x)] \cdot [(30 + x)/(120 + x)]$$

Setting $x/(120 + x) = [(40 + x)/(120 + x)] \cdot [(30 + x)/(120 + x)]$ implies that $x = 24$ black women need to be on the staff in order for A and B to be independent.

Table 2.5.1

	White	Black
Male	50	30
Female	40	

■

Comment Having shown that "Employee is female" and "Employee is black" are independent, does it follow that, say, "Employee is male" and "Employee is white" are independent? Yes. By virtue of the derivation in Example 2.5.2, the independence of events A and B implies the independence of events A^C and B^C (as well as A and B^C and A^C and B). It follows, then, that the $x = 24$ black women not only makes A and B independent, it also implies, more generally, that "race" and "sex" are independent.

Example 2.5.4 Suppose that two events, A and B, each having nonzero probability, are mutually exclusive. Are they also independent?

No. If A and B are mutually exclusive, then $P(A \cap B) = 0$. But $P(A) \cdot P(B) > 0$ (by assumption), so the equality spelled out in Definition 2.5.1 that characterizes independence is not met. ■

DEDUCING INDEPENDENCE

Sometimes the physical circumstances surrounding two events make it obvious that the occurrence (or nonoccurrence) of one has absolutely no influence or effect on the occurrence (or nonoccurrence) of the other. If that should be the case, then the two events will necessarily be *independent* in the sense of Definition 2.5.1.

Suppose a coin is tossed twice. Clearly, whatever happens on the first toss has no physical connection or influence on the outcome of the second. If A and B, then, are events defined on the second and first tosses, respectively, it would have to be the case that $P(A|B) = P(A|B^C) = P(A)$. For example, let A be the event that the second toss of a fair coin is a head, and let B be the event that the first toss of that coin is a tail. Then

$$\begin{aligned} P(A|B) &= P(\text{Head on second toss} \mid \text{tail on first toss}) \\ &= P(\text{Head on second toss}) = \frac{1}{2} \end{aligned}$$

Being able to infer that certain events are independent proves to be of enormous help in solving certain problems. The reason is that many events of interest are, in fact, intersections. If those events are independent, then the probability of that intersection reduces to a simple product (because of Definition 2.5.1)—that is, $P(A \cap B) = P(A) \cdot P(B)$. For the coin tosses just described,

$$\begin{aligned} P(A \cap B) &= P(\text{head on second toss} \cap \text{tail on first toss}) \\ &= P(A) \cdot P(B) \\ &= P(\text{head on second toss}) \cdot P(\text{tail on first toss}) \\ &= \frac{1}{2} \cdot \frac{1}{2} \\ &= \frac{1}{4} \end{aligned}$$

Example 2.5.5

Myra and Carlos are summer interns working as proofreaders for a local newspaper. Based on aptitude tests, Myra has a 50% chance of spotting a hyphenation error, while Carlos picks up on that same kind of mistake 80% of the time. Suppose the copy they are proofing contains a hyphenation error. What is the probability it goes undetected?

Let A and B be the events that Myra and Carlos, respectively, catch the mistake. By assumption, $P(A) = 0.50$ and $P(B) = 0.80$. What we are looking for is the probability of the complement of a union. That is,

$$\begin{aligned} P(\text{Error goes undetected}) &= 1 - P(\text{Error } \textit{is} \text{ detected}) \\ &= 1 - P(\text{Myra or Carlos or both see the mistake}) \\ &= 1 - P(A \cup B) \\ &= 1 - [P(A) + P(B) - P(A \cap B)] \quad (\text{from Theorem 2.3.6}) \end{aligned}$$

Since proofreaders invariably work by themselves, events A and B are necessarily independent, so $P(A \cap B)$ would reduce to the product $P(A) \cdot P(B)$. It follows that such an error would go unnoticed 10% of the time:

$$\begin{aligned} P(\text{Error goes undetected}) &= 1 - \{0.50 + 0.80 - (0.50)(0.80)\} = 1 - 0.90 \\ &= 0.10 \end{aligned}$$

■

Example 2.5.6

Suppose that one of the genes associated with the control of carbohydrate metabolism exhibits two alleles—a dominant W and a recessive w. If the probabilities of the WW, Ww, and ww genotypes in the present generation are p, q, and r, respectively, for both males and females, what are the chances that an individual in the *next* generation will be a ww?

Let A denote the event that an offspring receives a w allele from her father; let B denote the event that she receives the recessive allele from her mother. What we are looking for is $P(A \cap B)$.

According to the information given,

$$\begin{aligned} p &= P(\text{Parent has genotype WW}) = P(\text{WW}) \\ q &= P(\text{Parent has genotype Ww}) = P(\text{Ww}) \\ r &= P(\text{Parent has genotype ww}) = P(\text{ww}) \end{aligned}$$

If an offspring is equally likely to receive either of her parent's alleles, the probabilities of A and B can be computed using Theorem 2.4.1:

$$\begin{aligned} P(A) &= P(A \mid \text{WW})P(\text{WW}) + P(A \mid \text{Ww})P(\text{Ww}) + P(A \mid \text{ww})P(\text{ww}) \\ &= 0 \cdot p + \frac{1}{2} \cdot q + 1 \cdot r \\ &= r + \frac{q}{2} = P(B) \end{aligned}$$

Lacking any evidence to the contrary, there is every reason here to assume that A and B are independent events, in which case

$$\begin{aligned} P(A \cap B) &= P(\text{Offspring has genotype ww}) \\ &= P(A) \cdot P(B) \\ &= \left(r + \frac{q}{2}\right)^2 \end{aligned}$$

(This particular model for allele segregation, together with the independence assumption, is called *random Mendelian mating.*) ■

The last two examples focused on events for which an *a priori* assumption of independence was eminently reasonable. Sometimes, though, what seems eminently reasonable turns out to be surprisingly incorrect. The next example is a case in point.

Example 2.5.7

A crooked gambler has nine dice in her coat pocket. Three are fair and six are not. The biased ones are loaded in such a way that the probability of rolling a 6 is 1/2. She takes out one die at random and rolls it twice. Let A be the event "6 appears on first roll" and let B be the event "6 appears on second roll." Are A and B independent?

Our intuition here would probably answer "yes": How can two rolls of a die *not* be independent? For every dice problem we have encountered so far, they have been. But this is not a typical dice problem. Repeated throws of a die *do* qualify as independent events *if* the probabilities associated with the different faces are known. In this situation, though, those probabilities are not known and depend in a random way on which die the gambler draws from her pocket.

To see what effect not knowing which die is being tossed has on the relationship between A and B requires an application of Theorem 2.4.1. Let F and L denote the events "fair die is selected" and "loaded die is selected", respectively. Then

$$\begin{aligned} P(A \cap B) &= P(6 \text{ on first roll} \cap 6 \text{ on second roll}) \\ &= P(A \cap B \mid F)P(F) + P(A \cap B \mid L)P(L) \end{aligned}$$

Conditional on either F or L, A and B are independent, so

$$P(A \cap B) = (1/6)(1/6)(3/9) + (1/2)(1/2)(6/9) = 19/108$$

Similarly,

$$\begin{aligned} P(A) &= P(A \mid F)P(F) + P(A \mid L)P(L) \\ &= (1/6)(3/9) + (1/2)(6/9) = 7/18 = P(B) \end{aligned}$$

But note that

$$P(A \cap B) = 19/108 = 57/324 \neq P(A) \cdot P(B) = (7/18)(7/18) = 49/324$$

proving that A and B are *not* independent. ■

Questions

2.5.1. Suppose that $P(A \cap B) = 0.2$, $P(A) = 0.6$, and $P(B) = 0.5$.

(a) Are A and B mutually exclusive?

(b) Are A and B independent?

(c) Find $P(A^C \cup B^C)$.

2.5.2. Spike is not a terribly bright student. His chances of passing chemistry are 0.35; mathematics, 0.40; and both, 0.12. Are the events "Spike passes chemistry" and "Spike passes mathematics" independent? What is the probability that he fails both subjects?

2.5.3. Two fair dice are rolled. What is the probability that the number showing on one will be twice the number appearing on the other?

2.5.4. Emma and Josh have just gotten engaged. What is the probability that they have different blood types? Assume that blood types for both men and women are distributed in the general population according to the following proportions:

Blood Type	Proportion
A	40%
B	10%
AB	5%
O	45%

2.5.5. Dana and Cathy are playing tennis. The probability that Dana wins at least one out of two games is 0.3.

What is the probability that Dana wins at least one out of four?

2.5.6. Three points, X_1, X_2, and X_3, are chosen at random in the interval $(0, a)$. A second set of three points, Y_1, Y_2, and Y_3, are chosen at random in the interval $(0, b)$. Let A be the event that X_2 is between X_1 and X_3. Let B be the event that $Y_1 < Y_2 < Y_3$. Find $P(A \cap B)$.

2.5.7. Suppose that $P(A) = \frac{1}{4}$ and $P(B) = \frac{1}{8}$.

(a) What does $P(A \cup B)$ equal if

1. A and B are mutually exclusive?
2. A and B are independent?

(b) What does $P(A \mid B)$ equal if

1. A and B are mutually exclusive?
2. A and B are independent?

2.5.8. Suppose that events A, B, and C are independent.

(a) Use a Venn diagram to find an expression for $P(A \cup B \cup C)$ that does *not* make use of a complement.

(b) Find an expression for $P(A \cup B \cup C)$ that *does* make use of a complement.

2.5.9. A fair coin is tossed four times. What is the probability that the number of heads appearing on the first two tosses is equal to the number of heads appearing on the second two tosses?

2.5.10. Suppose that two cards are drawn simultaneously from a standard fifty-two-card poker deck. Let A be the event that both are either a jack, queen, king, or ace of hearts, and let B be the event that both are aces. Are A and B independent? (*Note:* There are 1326 equally likely ways to draw two cards from a poker deck.)

DEFINING THE INDEPENDENCE OF MORE THAN TWO EVENTS

It is not immediately obvious how to extend Definition 2.5.1 to, say, *three* events. To call A, B, and C independent, should we require that the probability of the three-way intersection factors into the product of the three original probabilities,

$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C) \tag{2.5.3}$$

or should we impose the definition we already have on the three *pairs* of events:

$$\begin{aligned} P(A \cap B) &= P(A) \cdot P(B) \\ P(B \cap C) &= P(B) \cdot P(C) \\ P(A \cap C) &= P(A) \cdot P(C) \end{aligned} \tag{2.5.4}$$

Actually, neither condition by itself is sufficient. If three events satisfy Equations 2.5.3 *and* 2.5.4, we will call them independent (or *mutually independent*), but Equation 2.5.3 does not imply Equation 2.5.4, nor does Equation 2.5.4 imply Equation 2.5.3 (see Questions 2.5.11 and 2.5.12).

More generally, the independence of n events requires that the probabilities of all possible intersections equal the products of all the corresponding individual probabilities. Definition 2.5.2 states the result formally. Analogous to what was true in the case of *two* events, the practical applications of Definition 2.5.2 arise when n events are mutually independent, and we can calculate $P(A_1 \cap A_2 \cap \cdots \cap A_n)$ by computing the product $P(A_1) \cdot P(A_2) \cdots P(A_n)$.

Definition 2.5.2
Events $A_1, A_2, \ldots, A_n$ are said to be *independent* if for every set of indices i_1, $i_2, \ldots, i_k$ between 1 and n, inclusive,

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

Example 2.5.8

An insurance company plans to assess its future liabilities by sampling the records of its current policyholders. A pilot study has turned up three clients—one living in Alaska, one in Missouri, and one in Vermont—whose estimated chances of surviving to the year 2020 are 0.7, 0.9, and 0.3, respectively. What is the probability that by the end of 2019 the company will have had to pay death benefits to exactly one of the three?

Let A_1 be the event "Alaska client survives through 2019." Define A_2 and A_3 analogously for the Missouri client and Vermont client, respectively. Then the event E: "Exactly one dies" can be written as the union of three intersections:

$$E = \left(A_1 \cap A_2 \cap A_3^C\right) \cup \left(A_1 \cap A_2^C \cap A_3\right) \cup \left(A_1^C \cap A_2 \cap A_3\right)$$

Since each of the intersections is mutually exclusive of the other two,

$$P(E) = P\left(A_1 \cap A_2 \cap A_3^C\right) + P\left(A_1 \cap A_2^C \cap A_3\right) + P\left(A_1^C \cap A_2 \cap A_3\right)$$

Furthermore, there is no reason to believe that for all practical purposes the fates of the three are not independent. That being the case, each of the intersection probabilities reduces to a product, and we can write

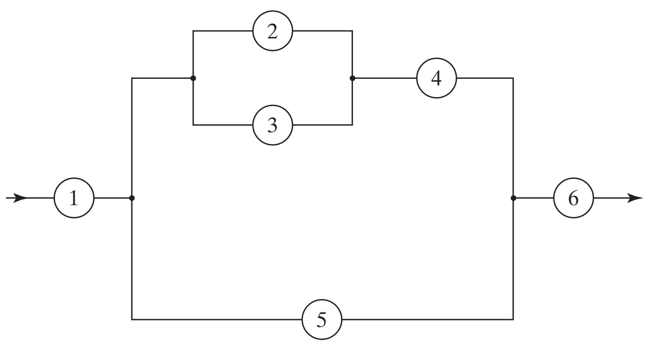
$$\begin{aligned} P(E) &= P(A_1) \cdot P(A_2) \cdot P\left(A_3^C\right) + P(A_1) \cdot P\left(A_2^C\right) \cdot P(A_3) + P\left(A_1^C\right) \cdot P(A_2) \cdot P(A_3) \\ &= (0.7)(0.9)(0.7) + (0.7)(0.1)(0.3) + (0.3)(0.9)(0.3) \\ &= 0.543 \end{aligned}$$

■

Example 2.5.9

Protocol for making financial decisions in a certain corporation follows the "circuit" pictured in Figure 2.5.1. Any budget is first screened by 1. If he approves it, the plan is forwarded to 2, 3, and 5. If either 2 or 3 concurs, it goes to 4. If either 4 or 5 says "yes," it moves on to 6 for a final reading. Only if 6 is also in agreement does the proposal pass. Suppose that 1, 5, and 6 each has a 50% chance of saying "yes," whereas 2, 3, and 4 will each concur with a probability of 0.70. If everyone comes to a decision independently, what is the probability that a budget will pass?

Figure 2.5.1



Probabilities of this sort are calculated by reducing the circuit to its component unions and intersections. Moreover, if all decisions are made independently, which is the case here, then every intersection becomes a product.

Let A_i be the event that person i approves the budget, $i = 1, 2, \ldots, 6$. Looking at Figure 2.5.1, we see that

$$\begin{aligned} P(\text{Budget passes}) &= P(A_1 \cap \{[(A_2 \cup A_3) \cap A_4] \cup A_5\} \cap A_6) \\ &= P(A_1)P\{[(A_2 \cup A_3) \cap A_4] \cup A_5\}P(A_6) \end{aligned}$$

By assumption, $P(A_1) = 0.5$, $P(A_2) = 0.7$, $P(A_3) = 0.7$, $P(A_4) = 0.7$, $P(A_5) = 0.5$, and $P(A_6) = 0.5$, so

$$\begin{aligned} P\{[(A_2 \cup A_3) \cap A_4]\} &= [P(A_2) + P(A_3) - P(A_2)P(A_3)]P(A_4) \\ &= [0.7 + 0.7 - (0.7)(0.7)](0.7) \\ &= 0.637 \end{aligned}$$

Therefore,

$$\begin{aligned} P(\text{Budget passes}) &= (0.5)\{0.637 + 0.5 - (0.637)(0.5)\}(0.5) \\ &= 0.205 \end{aligned}$$

■

REPEATED INDEPENDENT EVENTS

We have already seen several examples where the event of interest was actually an intersection of independent simpler events (in which case the probability of the intersection reduced to a product). There is a special case of that basic scenario that deserves special mention because it applies to numerous real-world situations. If the events making up the intersection all arise from the same physical circumstances and assumptions (i.e., they represent repetitions of the same experiment), they are referred to as *repeated independent trials*. The number of such trials may be finite or infinite.

Example 2.5.10 Suppose the string of Christmas tree lights you just bought has twenty-four bulbs wired in series. If each bulb has a 99.9% chance of "working" the first time current is applied, what is the probability that the string itself will *not* work?

Let A_i be the event that the ith bulb fails, $i = 1, 2, \ldots, 24$. Then

$$\begin{aligned} P(\text{String fails}) &= P(\text{At least one bulb fails}) \\ &= P(A_1 \cup A_2 \cup \cdots \cup A_{24}) \\ &= 1 - P(\text{String works}) \\ &= 1 - P(\text{All twenty-four bulbs work}) \\ &= 1 - P\left(A_1^C \cap A_2^C \cap \cdots \cap A_{24}^C\right) \end{aligned}$$

If we assume that bulb failures are independent events,

$$P(\text{String fails}) = 1 - P\left(A_1^C\right)P\left(A_2^C\right) \cdots P\left(A_{24}^C\right)$$

Moreover, since all the bulbs are presumably manufactured the same way, $P(A_i^C)$ is the same for all i, so

$$\begin{aligned} P(\text{String fails}) &= 1 - \left[P\left(A_i^C\right)\right]^{24} \\ &= 1 - (0.999)^{24} \\ &= 1 - 0.98 \\ &= 0.02 \end{aligned}$$

The chances are one in fifty, in other words, that the string will not work the first time current is applied. ∎

Example 2.5.11 During the 1978 baseball season, Pete Rose of the Cincinnati Reds set a National League record by hitting safely in forty-four consecutive games. Assume that Rose was a .300 hitter and that he came to bat four times each game. If each at-bat is assumed to be an independent event, what probability might reasonably be associated with a hitting streak of that length?

For this problem we need to invoke the repeated independent trials model *twice*—once for the four at-bats making up a game and a second time for the forty-four games making up the streak. Let A_i denote the event "Rose hit safely in ith game," $i = 1, 2, \ldots, 44$. Then

$$\begin{aligned} P(\text{Rose hit safely in forty-four consecutive games}) &= P(A_1 \cap A_2 \cap \cdots \cap A_{44}) \\ &= P(A_1) \cdot P(A_2) \cdot \cdots \cdot P(A_{44}) \end{aligned} \tag{2.5.5}$$

Since all the $P(A_i)$'s are equal, we can further simplify Equation 2.5.5 by writing

$$P(\text{Rose hit safely in forty-four consecutive games}) = [P(A_1)]^{44}$$

To calculate $P(A_1)$ we should focus on the *complement* of A_1. Specifically,

$$\begin{aligned} P(A_1) &= 1 - P\left(A_1^C\right) \\ &= 1 - P(\text{Rose did } \textit{not} \text{ hit safely in game 1}) \\ &= 1 - P(\text{Rose made four outs}) \\ &= 1 - (0.700)^4 \quad (\text{Why?}) \\ &= 0.76 \end{aligned}$$

Therefore, the probability of a .300 hitter putting together a forty-four-game streak (during a given set of forty-four games) is 0.0000057:

$$\begin{aligned} P(\text{Rose hit safely in forty-four consecutive games}) &= (0.76)^{44} \\ &= 0.0000057 \end{aligned}$$ ∎

Comment The analysis described here has the basic "structure" of a repeated independent trials problem, but the assumptions that the latter makes are not entirely satisfied by the data. Each at-bat, for example, is not really a repetition of the same experiment, nor is $P(A_i)$ the same for all i. Rose would obviously have had different probabilities of getting a hit against different pitchers. Moreover, although "four" was probably the typical number of official at-bats that he had during a game, there would certainly have been many instances where he had either fewer or more. Clearly, the day-to-day deviations from the assumed model would have sometimes been in Rose's favor, sometimes not. Over the course of the hitting streak, the net effect of those deviations would not be expected to have much effect on the 0.0000057 probability.

Example 2.5.12 In the game of craps, one of the ways a player can win is by rolling (with two dice) one of the sums 4, 5, 6, 8, 9, or 10, and then rolling that sum again before rolling a sum of 7. For example, the sequence of sums 6, 5, 8, 8, 6 would result in the player winning on his fifth roll. In gambling parlance, "6" is the player's "point," and he "made his point." On the other hand, the sequence of sums 8, 4, 10, 7 would result in the player losing on his fourth roll: his point was an 8, but he rolled a sum of 7 before he rolled a second 8. What is the probability that a player wins with a point of 10?

Table 2.5.2 shows some of the ways a player can make a point of 10. Each sequence, of course, is an intersection of independent events, so its probability becomes

Table 2.5.2

Sequence of Rolls	Probability
(10, 10)	(3/36)(3/36)
(10, no 10 or 7, 10)	(3/36)(27/36)(3/36)
(10, no 10 or 7, no 10 or 7, 10)	(3/36)(27/36)(27/36)(3/36)
⋮	⋮

a product. The event "Player wins with a point of 10" is then the union of all the sequences that could have been listed in the first column. Since all those sequences are mutually exclusive, the probability of winning with a point of 10 reduces to the sum of an infinite number of products:

$$\begin{aligned} P(\text{Player wins with a point of 10}) &= \frac{3}{36}\cdot\frac{3}{36}+\frac{3}{36}\cdot\frac{27}{36}\cdot\frac{3}{36} \\ &\quad+\frac{3}{36}\cdot\frac{27}{36}\cdot\frac{27}{36}\cdot\frac{3}{36}+\cdots \\ &= \frac{3}{36}\cdot\frac{3}{36}\sum_{k=0}^{\infty}\left(\frac{27}{36}\right)^k \end{aligned} \tag{2.5.6}$$

Recall from algebra that if $0 < r < 1$,

$$\sum_{k=0}^{\infty} r^k = 1/(1-r)$$

Applying the formula for the sum of a geometric series to Equation 2.5.6 shows that the probability of winning at craps with a point of 10 is $\frac{1}{36}$:

$$\begin{aligned} P(\text{Player wins with a point of 10}) &= \frac{3}{36}\cdot\frac{3}{36}\cdot\frac{1}{(1-\frac{27}{36})} \\ &= \frac{1}{36} \end{aligned}$$

Table 2.5.3 shows the probabilities of a person "making" each of the possible six points—4, 5, 6, 8, 9, and 10. According to the rules of craps, a player wins by either

Table 2.5.3

Point	P (makes point)
4	1/36
5	16/360
6	25/396
8	25/396
9	16/360
10	1/36

(1) getting a sum of 7 or 11 on the first roll or (2) getting a 4, 5, 6, 8, 9, or 10 on the first roll and making the point. But $P(\text{sum} = 7) = 6/36$ and $P(\text{sum} = 11) = 2/36$, so

$$\begin{aligned} P(\text{Player wins}) &= \frac{6}{36} + \frac{2}{36} + \frac{1}{36} + \frac{16}{360} + \frac{25}{396} + \frac{25}{396} + \frac{16}{360} + \frac{1}{36} \\ &= 0.493 \end{aligned}$$

As even-money games go, craps is relatively fair—the probability of the shooter winning is not much less than 0.500. ■

Example 2.5.13 A transmitter is sending a binary code (+ and − signals) that must pass through three relay signals before being sent on to the receiver (see Figure 2.5.2). At each relay station, there is a 25% chance that the signal will be reversed—that is

$$\begin{aligned} &P(+ \text{ is sent by relay } i \mid - \text{ is received by relay } i) \\ &\quad = P(- \text{ is sent by relay } i \mid + \text{ is received by relay } i) \\ &\quad = 1/4, \qquad i = 1, 2, 3 \end{aligned}$$

Suppose + symbols make up 60% of the message being sent. If the signal + is received, what is the probability a + was sent?

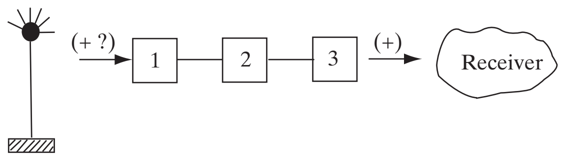


Figure 2.5.2

This is basically a Bayes' theorem (Theorem 2.4.2) problem, but the three relay stations introduce a complex mechanism for transmission error. Let A be the event "+ is transmitted from tower" and B be the event "+ is received from relay 3." Then

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^C)P(A^C)}$$

Notice that a + can be received from relay 3 given that a + was initially sent from the tower if either (1) all relay stations function properly or (2) any *two* of the stations make transmission errors. Table 2.5.4 shows the four mutually exclusive ways (1) and (2) can happen. The probabilities associated with the message transmissions at each relay station are shown in parentheses. Assuming the relay station outputs are independent events, the probability of an entire transmission sequence is simply the product of the probabilities in parentheses in any given row. These overall probabilities are listed in the last column; their sum, 36/64, is $P(B|A)$. By a similar analysis, we can show that

$$P(B|A^C) = P(+ \text{ is received from relay 3} \mid - \text{ is transmitted from tower}) = 28/64$$

Finally, since $P(A) = 0.6$ and $P(A^C) = 0.4$, the conditional probability we are looking for is 0.66:

$$P(A|B) = \frac{\left(\frac{36}{64}\right)(0.6)}{\left(\frac{36}{64}\right)(0.6) + \left(\frac{28}{64}\right)(0.4)} = 0.66$$ ■

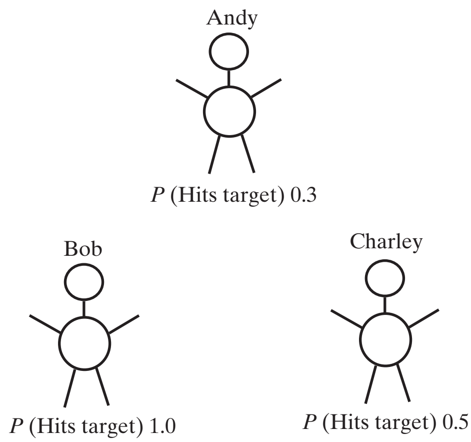
Table 2.5.4

	Signal Transmitted by			
Tower	Relay 1	Relay 2	Relay 3	Probability
+	+(3/4)	−(1/4)	+(1/4)	3/64
+	−(1/4)	−(3/4)	+(1/4)	3/64
+	−(1/4)	+(1/4)	+(3/4)	3/64
+	+(3/4)	+(3/4)	+(3/4)	27/64
				36/64

Example 2.5.14

Andy, Bob, and Charley have gotten into a disagreement over a female acquaintance, Donna, and decide to settle their dispute with a three-cornered pistol duel. Of the three, Andy is the worst shot, hitting his target only 30% of the time. Charley, a little better, is on-target 50% of the time, while Bob never misses (see Figure 2.5.3). The rules they agree to are simple: They are to fire at the targets of their choice in succession, and cyclically, in the order Andy, Bob, Charley, and so on, until only one of them is left standing. On each "turn," they get only one shot. If a combatant is hit, he no longer participates, either as a target or as a shooter.

Figure 2.5.3



As Andy loads his revolver, he mulls over his options (his objective is clear—to maximize his probability of survival). According to the rule he can shoot either Bob or Charley, but he quickly rules out shooting at the latter because it would be counterproductive to his future well-being: If he shot at Charley and had the misfortune of hitting him, it would be Bob's turn, and Bob would have no recourse but to shoot at Andy. From Andy's point of view, this would be a decidedly grim turn of events, since Bob never misses. Clearly, Andy's only option is to shoot at Bob. This leaves two scenarios: (1) He shoots at Bob and hits him, or (2) he shoots at Bob and misses.

Consider the first possibility. If Andy hits Bob, Charley will proceed to shoot at Andy, Andy will shoot back at Charley, and so on, until one of them hits the other. Let CH_i and CM_i denote the events "Charley hits Andy with the ith shot" and "Charley misses Andy with the ith shot," respectively. Define AH_i and AM_i analogously. Then Andy's chances of survival (given that he has killed Bob) reduce to a countably infinite union of intersections:

$$\begin{aligned}P(\text{Andy survives}) = P[&(CM_1 \cap AH_1) \cup (CM_1 \cap AM_1 \cap CM_2 \cap AH_2) \\ &\cup (CM_1 \cap AM_1 \cap CM_2 \cap AM_2 \cap CM_3 \cap AH_3) \cup \cdots]\end{aligned}$$

Note that each intersection is mutually exclusive of all of the others and that its component events are independent. Therefore,

$$\begin{aligned}
P(\text{Andy survives}) &= P(CM_1)P(AH_1) + P(CM_1)P(AM_1)P(CM_2)P(AH_2) \\
&\quad + P(CM_1)P(AM_1)P(CM_2)P(AM_2)P(CM_3)P(AH_3) + \cdots \\
&= (0.5)(0.3) + (0.5)(0.7)(0.5)(0.3) + (0.5)(0.7)(0.5)(0.7)(0.5)(0.3) + \cdots \\
&= (0.5)(0.3)\sum_{k=0}^{\infty}(0.35)^k \\
&= (0.15)\left(\frac{1}{1-0.35}\right) = \frac{3}{13}
\end{aligned}$$

Now consider the second scenario. If Andy shoots at Bob and misses, Bob will undoubtedly shoot and hit Charley, since Charley is the more dangerous adversary. Then it will be Andy's turn again. Whether he would see another tomorrow would depend on his ability to make that very next shot count. Specifically,

$$P(\text{Andy survives}) = P(\text{Andy hits Bob on second turn}) = 3/10$$

But $\frac{3}{10} > \frac{3}{13}$, so Andy is better off *not* hitting Bob with his first shot. And because we have already argued that it would be foolhardy for Andy to shoot at Charley, Andy's optimal strategy is clear—deliberately miss both Bob and Charley with the first shot. ■

Example 2.5.15

Scientists estimate that the Earth's atmosphere contains on the order of 10^{44} molecules of which roughly 78% are Nitrogen (N_2), 21% Oxygen (O_2), and 1% Argon (A). Some 2×10^{22} of those molecules make up each and every breath you take. Rooted in those not particularly interesting facts is a bizarre question, one that has an equally bizarre answer.

On March 15, 44 B.C., Julius Caesar was assassinated by a group of Roman senators, a mutiny led by one of his dearest friends, Marcus Brutus. In Shakespeare's play describing the attack, the dying Caesar calls out his friend's treachery with the famous lament, "Et tu, Brute." In that final breath, Caesar presumably exhaled 2×10^{22} molecules.

So, here is the question, framed in the context of a typical balls-in-boxes problem. Suppose those last 2×10^{22} molecules that Caesar exhaled could all be colored red; and suppose the remaining $10^{44} - 2 \times 10^{22}$ molecules in the atmosphere could all be colored blue. Furthermore, suppose the entire set of red and blue molecules remains unchanged over time, but the box (that is, the atmosphere) is stirred and shaken every day for the next 2060 years. By 2016, then, we can assume that Caesar's last breath (of red molecules) has become randomly scattered throughout Earth's atmosphere (of blue molecules). Given those assumptions, what is the probability that *in your next breath* there will be at least one molecule from Caesar's *last breath*?

At first blush, everyone's intuition would dismiss the question as absurd—using a sample as small as a single breath to "recapture" something from another breath that could be anywhere in the atmosphere would seem to be analogous to a blind person searching for an infinitesimally small needle in a Godzilla-sized haystack. For all practical purposes, the answer surely must be 0. Not so.

Imagine taking a random sample of 2×10^{22} molecules (your next breath) *one-at-a-time* from the 10^{44} molecules in the atmosphere (see Figure 2.5.4). Let A_i, $i = 1, 2, \ldots, 2 \times 10^{22}$ be the event that the ith molecule taken in your next breath is *not* a

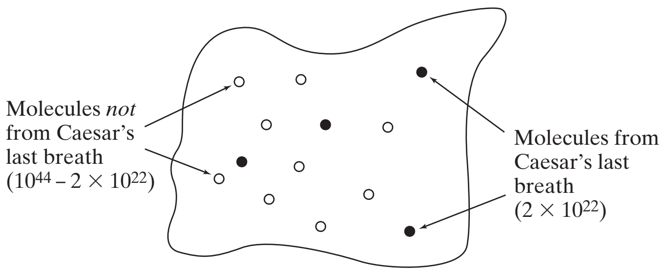


Figure 2.5.4

"Caesar" molecule, and let A denote the event that your last breath ultimately *does* contain something from Caesar's last breath. Then

$$P(A) = 1 - P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_{2\times 10^{22}})$$

Clearly,

$$P(A_1) = \left(10^{44} - 2 \times 10^{22}\right)/10^{44}$$
$$P(A_2|A_1) = \left(10^{44} - 2 \times 10^{22} - 1\right)/\left(10^{44} - 1\right)$$
$$P(A_3|A_1 \cap A_2) = \left(10^{44} - 2 \times 10^{22} - 2\right)/\left(10^{44} - 2\right), \text{ and so on.}$$

Also, from Equation 2.4.3,

$$P(A) = 1 - P(A_1)\,P(A_2|A_1)\,P(A_3|A_1 \cap A_2)\ldots P(A_{2\times 10^{22}}|A_1 \cap A_2 \cap \ldots A_{2\times 10^{22}-1})$$

Because of the huge disparity, though, between 10^{44} and 2×10^{22}, all the conditional probabilities are essentially equal to $P(A^1)$, which implies that

$$P(A) \doteq 1 - [P(A_1)]^{2\times 10^{22}} = 1 - \left[(10^{44} - 2 \times 10^{22})/10^{44}\right]^{2\times 10^{22}}$$
$$= 1 - \left[(1 - 2/10^{22})\right]^{2\times 10^{22}}$$

Recall from calculus that if x is a very small number, $1 - x \doteq e^{-x}$. Therefore,

$$P(A) \doteq 1 - (e^{-2/10^{22}})^{2\times 10^{22}} = 1 - e^{-4} = 0.98 \tag{2.5.7}$$

Equation 2.5.7 is a shocker to say the least: Given the assumptions that were made, it shows that the probability your next breath contains something from Caesar's last breath is almost a certainty. How can something so unbelievable have such a high probability of happening? The answer in a word is *persistence*. Think of your next breath as a raffle with anything from Caesar's last breath being the top prize. What makes this "game" different than a weekly LOTTO drawing where you might have purchased a half-dozen tickets is that here your next breath, in effect, has purchased 20,000,000,000,000,000,000,000 tickets. Moreover, there is not just *one* jackpot, there are *20,000,000,000,000,000,000,000* jackpots.

The great Roman poet and philosopher Ovid was born a year before Caesar was murdered. Among his voluminous writings was an interesting reflection on the nature of probability, one suggesting that he might not have been surprised by the answer given in Equation 2.5.7. "Chance," he wrote, "is always powerful. Let your hook be always cast; in the pool where you least expect it, there will be a fish." ■

Questions

2.5.11. Suppose that two fair dice (one red and one green) are rolled. Define the events

A: a 1 or a 2 shows on the red die
B: a 3, 4, or 5 shows on the green die
C: the dice total is 4, 11, or 12

Show that these events satisfy Equation 2.5.3 but not Equation 2.5.4.

2.5.12. A roulette wheel has thirty-six numbers colored red or black according to the pattern indicated below:

Roulette wheel pattern

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
R	R	R	R	R	B	B	B	B	R	R	R	R	B	B	B	B	B
36	35	34	33	32	31	30	29	28	27	26	25	24	23	22	21	20	19

Define the events

A: red number appears
B: even number appears
C: number is less than or equal to 18

Show that these events satisfy Equation 2.5.4 but not Equation 2.5.3.

2.5.13. How many probability equations need to be verified to establish the mutual independence of *four* events?

2.5.14. In a roll of a pair of fair dice (one red and one green), let A be the event the red die shows a 3, 4, or 5; let B be the event the green die shows a 1 or a 2; and let C be the event the dice total is 7. Show that A, B, and C are independent.

2.5.15. In a roll of a pair of fair dice (one red and one green), let A be the event of an odd number on the red die, let B be the event of an odd number on the green die, and let C be the event that the sum is odd. Show that any pair of these events is independent but that A, B, and C are not mutually independent.

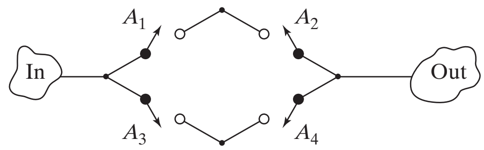
2.5.16. On her way to work, a commuter encounters four traffic signals. Assume that the distance between each of the four is sufficiently great that her probability of getting a green light at any intersection is independent of what happened at any previous intersection. The first two lights are green for forty seconds of each minute; the last two, for thirty seconds of each minute. What is the probability that the commuter has to stop at least three times?

2.5.17. School board officials are debating whether to require all high school seniors to take a proficiency exam before graduating. A student passing all three parts (mathematics, language skills, and general knowledge) would be awarded a diploma; otherwise, he or she would receive only a certificate of attendance. A practice test given to this year's ninety-five hundred seniors resulted in the following numbers of failures:

Subject Area	Number of Students Failing
Mathematics	3325
Language skills	1900
General knowledge	1425

If "Student fails mathematics," "Student fails language skills," and "Student fails general knowledge" are independent events, what proportion of next year's seniors can be expected to fail to qualify for a diploma? Does independence seem a reasonable assumption in this situation?

2.5.18. Consider the following four-switch circuit:



If all switches operate independently and P(Switch closes) $= p$, what is the probability the circuit is completed?

2.5.19. A fast-food chain is running a new promotion. For each purchase, a customer is given a game card that may win \$10. The company claims that the probability of a person winning at least once in five tries is 0.32. What is the probability that a customer wins \$10 on his or her first purchase?

2.5.20. Players A, B, and C toss a fair coin in order. The first to throw a head wins. What are their respective chances of winning?

2.5.21. In a certain developing nation, statistics show that only two out of ten children born in the early 1980s reached the age of twenty-one. If the same mortality rate is operative over the next generation, how many children does a woman need to bear if she wants to have at least a 75% probability that at least one of her offspring survives to adulthood?

2.5.22. According to an advertising study, 15% of television viewers who have seen a certain automobile commercial can correctly identify the actor who does the voice-over. Suppose that ten such people are watching TV and the commercial comes on. What is the probability that at least one of them will be able to name the actor? What is the probability that exactly one will be able to name the actor?

2.5.23. A fair die is rolled and then n fair coins are tossed, where n is the number showing on the die. What is the probability that no heads appear?

2.5.24. Each of m urns contains three red chips and four white chips. A total of r samples with replacement are taken from each urn. What is the probability that at least one red chip is drawn from at least one urn?

2.5.25. If two fair dice are tossed, what is the smallest number of throws, n, for which the probability of getting at least one double 6 exceeds 0.5? (*Note:* This was one of the first problems that de Méré communicated to Pascal in 1654.)

2.5.26. A pair of fair dice are rolled until the first sum of 8 appears. What is the probability that a sum of 7 does not precede that first sum of 8?

2.5.27. An urn contains w white chips, b black chips, and r red chips. The chips are drawn out at random, one at a time, with replacement. What is the probability that a white appears before a red?

2.5.28. A Coast Guard dispatcher receives an SOS from a ship that has run aground off the shore of a small island. Before the captain can relay her exact position, though, her radio goes dead. The dispatcher has n helicopter crews he can send out to conduct a search. He suspects the ship is somewhere either south in area I (with probability p) or north in area II (with probability $1 - p$). Each of the n rescue parties is equally competent and has probability r of locating the ship given it has run aground in the sector being searched. How should the dispatcher deploy the helicopter crews to maximize the probability that one of them will find the missing ship? (*Hint:* Assume that m search crews are sent to area I and $n - m$ are sent to area II. Let B denote the event that the ship is found, let A_1 be the event that the ship is in area I, and let A_2 be the event that the ship is in area II. Use Theorem 2.4.1 to get an expression for $P(B)$; then differentiate with respect to m.)

2.5.29. A computer is instructed to generate a random sequence using the digits 0 through 9; repetitions are permissible. What is the shortest length the sequence can be and still have at least a 70% probability of containing at least one 4?

2.5.30. A box contains a two-headed coin and eight fair coins. One coin is drawn at random and tossed n times. Suppose all n tosses come up heads. Show that the limit of the probability that the coin is fair is 0 as n goes to infinity.

2.5.31. Stanley's statistics seminar is graded on a Pass/Fail basis. At the end of the semester each student is given the option of taking either a two-question exam (Final A) or a three-question exam (Final B). To pass the course, students must answer at least one question correctly on whichever exam they choose. The professor estimates that a typical student has a 45% chance of correctly answering each of the two questions on Final A and a 30% chance of correctly answering each of the three questions on Final B. Which exam should Stanley choose? Answer the question two different ways.

2.5.32. What is the smallest number of switches wired in parallel that will give a probability of at least 0.98 that a circuit will be completed? Assume that each switch operates independently and will function properly 60% of the time.

2.6 Combinatorics

Combinatorics is a time-honored branch of mathematics concerned with counting, arranging, and ordering. While blessed with a wealth of early contributors (there are references to combinatorial problems in the Old Testament), its emergence as a separate discipline is often credited to the German mathematician and philosopher Gottfried Wilhelm Leibniz (1646–1716), whose 1666 treatise, *Dissertatio de arte combinatoria*, was perhaps the first monograph written on the subject (114).

Applications of combinatorics are rich in both diversity and number. Users range from the molecular biologist trying to determine how many ways genes can be positioned along a chromosome, to a computer scientist studying queuing priorities, to a psychologist modeling the way we learn, to a weekend poker player wondering whether he should draw to a straight, or a flush, or a full house. Surprisingly enough, despite the considerable differences that seem to distinguish one question from another, solutions to all of these questions are rooted in the same set of four basic theorems and rules.

COUNTING ORDERED SEQUENCES: THE MULTIPLICATION RULE

More often than not, the relevant "outcomes" in a combinatorial problem are ordered sequences. If two dice are rolled, for example, the outcome (4, 5)—that is, the first die comes up 4 and the second die comes up 5—is an ordered sequence of length two. The number of such sequences is calculated by using the most fundamental result in combinatorics, the *multiplication rule*.

Multiplication Rule *If operation A can be performed in m different ways and operation B in n different ways, the sequence (operation A, operation B) can be performed in $m \cdot n$ different ways.*

Proof At the risk of belaboring the obvious, we can verify the multiplication rule by considering a *tree* diagram (see Figure 2.6.1). Since each version of A can be followed by any of n versions of B, and there are m of the former, the total number of "A, B" sequences that can be pieced together is obviously the product $m \cdot n$. ∎

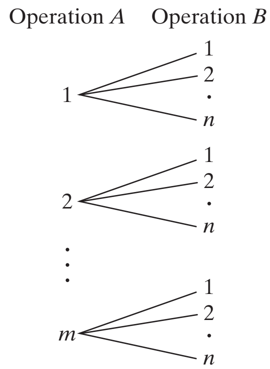


Figure 2.6.1

Corollary 2.6.1	*If operation A_i, $i = 1, 2, \ldots, k$, can be performed in n_i ways, $i = 1, 2, \ldots, k$, respectively, then the ordered sequence (operation A_1, operation A_2, ..., operation A_k) can be performed in $n_1 \cdot n_2 \cdot \cdots \cdot n_k$ ways.*

Example 2.6.1 The combination lock on a briefcase has two dials, each marked off with sixteen notches (see Figure 2.6.2). To open the case, a person first turns the left dial in a certain direction for two revolutions and then stops on a particular mark. The right dial is set in a similar fashion, after having been turned in a certain direction for two revolutions. How many different settings are possible?

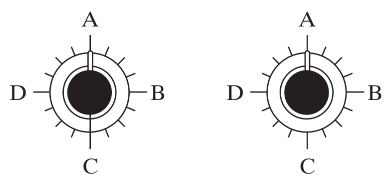


Figure 2.6.2

In the terminology of the multiplication rule, opening the briefcase corresponds to the four-step sequence (A_1, A_2, A_3, A_4) detailed in Table 2.6.1. Applying the previous corollary, we see that one thousand twenty-four different settings are possible:

$$\begin{aligned}\text{number of different settings} &= n_1 \cdot n_2 \cdot n_3 \cdot n_4 \\ &= 2 \cdot 16 \cdot 2 \cdot 16 \\ &= 1024\end{aligned}$$

Table 2.6.1

Operation	Purpose	Number of Options
$\mathcal{A}_1$	Rotating the left dial in a particular direction	2
$\mathcal{A}_2$	Choosing an endpoint for the left dial	16
$\mathcal{A}_3$	Rotating the right dial in a particular direction	2
$\mathcal{A}_4$	Choosing an endpoint for the right dial	16

Comment The number of dials, as opposed to the number of notches on each dial, is the critical factor in determining how many different settings are possible. A two-dial lock, for example, where each dial has twenty notches, gives rise to only $2 \cdot 20 \cdot 2 \cdot 20 = 1600$ settings. If those forty notches, though, are distributed among *four* dials (ten to each dial), the number of different settings increases a hundredfold to 160,000 $(= 2 \cdot 10 \cdot 2 \cdot 10 \cdot 2 \cdot 10 \cdot 2 \cdot 10)$. ■

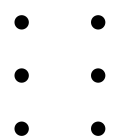
Example 2.6.2 Alphonse Bertillon, a nineteenth-century French criminologist, developed an identification system based on eleven anatomical variables (height, head width, ear length, etc.) that presumably remain essentially unchanged during an individual's adult life. The range of each variable was divided into three subintervals: small, medium, and large. A person's *Bertillon configuration* is an ordered sequence of eleven letters, say,

$$s, s, m, m, l, s, l, s, s, m, s$$

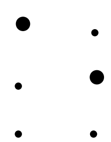
where a letter indicates the individual's "size" relative to a particular variable. How populated does a city have to be before it can be guaranteed that at least two citizens will have the same Bertillon configuration?

Viewed as an ordered sequence, a Bertillon configuration is an eleven-step classification system, where three options are available at each step. By the multiplication rule, a total of 3^{11}, or 177,147, distinct sequences are possible. Therefore, any city with at least 177,148 adults would necessarily have at least two residents with the same pattern. (The limited number of possibilities generated by the configuration's variables proved to be one of its major weaknesses. Still, it was widely used in Europe for criminal identification before the development of fingerprinting.) ■

Example 2.6.3 In 1824 Louis Braille invented what would eventually become the standard alphabet for the blind. Based on an earlier form of "night writing" used by the French army for reading battlefield communiqués in the dark, Braille's system replaced each written character with a six-dot matrix:



where certain dots were raised, the choice depending on the character being transcribed. The letter e, for example, has two raised dots and is written



Punctuation marks, common words, suffixes, and so on, also have specified dot patterns. In all, how many different characters can be enciphered in Braille?

Think of the dots as six distinct operations, numbered 1 to 6 (see Figure 2.6.3). In forming a Braille letter, we have two options for each dot: We can raise it or *not* raise it. The letter *e*, for example, corresponds to the six-step sequence (raise, do not raise, do not raise, do not raise, raise, do not raise). The number of such sequences, with $k = 6$ and $n_1 = n_2 = \cdots = n_6 = 2$, is 2^6, or 64. One of those sixty-four configurations, though, has *no* raised dots, making it of no use to a blind person. Figure 2.6.4 shows the entire sixty-three-character Braille alphabet. ■

Figure 2.6.3

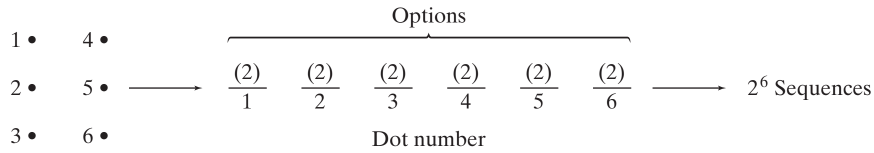
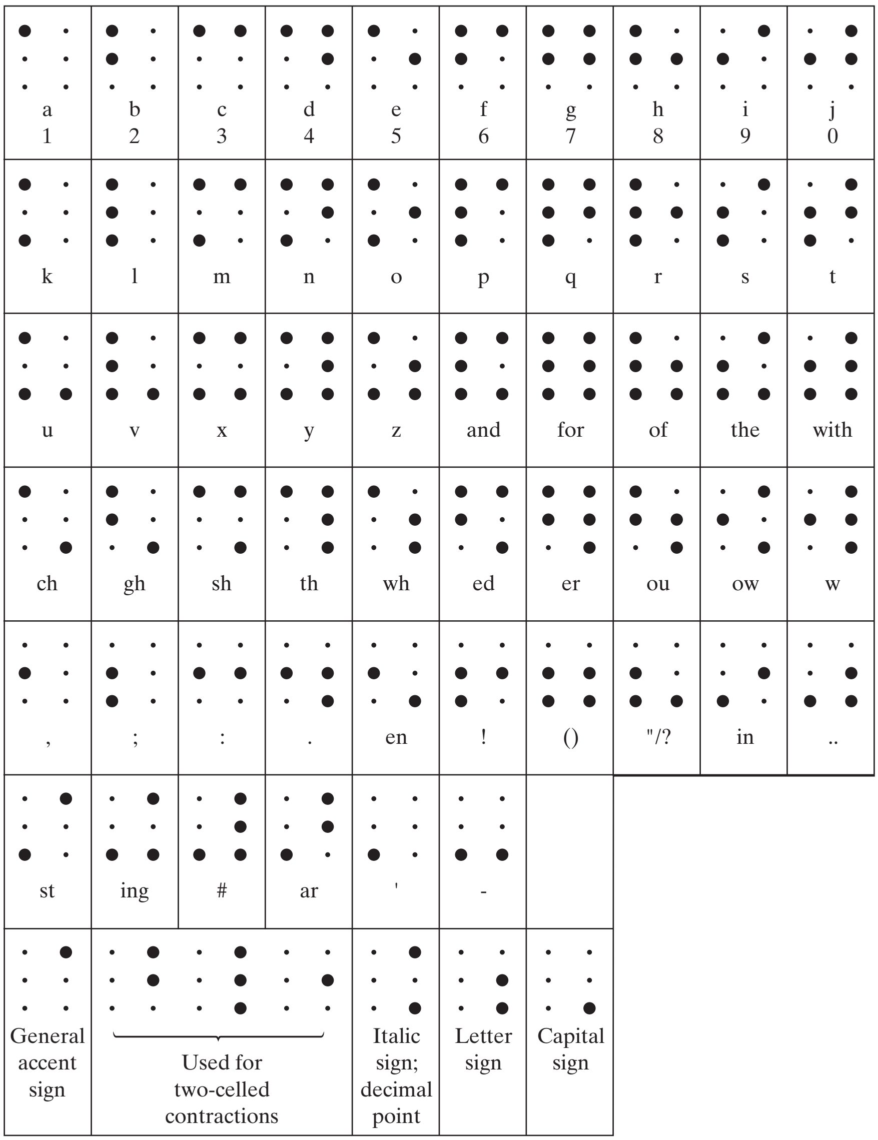


Figure 2.6.4



Example 2.6.4 The annual NCAA ("March Madness") basketball tournament starts with a field of sixty-four teams. After six rounds of play, the squad that remains unbeaten is declared the national champion. How many different configurations of winners and losers are possible, starting with the first round? Assume that the initial pairing of the sixty-four invited teams into thirty-two first-round matches has already been done.

Counting the number of ways a tournament of this sort can play out is an exercise in applying the multiplication rule twice. Notice, first, that the thirty-two first-round games can be decided in 2^{32} ways. Similarly, the resulting sixteen second-round games can generate 2^{16} different winners, and so on. Overall, the tournament can be pictured as a six-step sequence, where the number of possible outcomes at the six steps are 2^{32}, 2^{16}, 2^8, 2^4, 2^2, and 2^1, respectively. It follows that the number of possible tournaments (not all of which, of course, would be equally likely!) is the product $2^{32} \cdot 2^{16} \cdot 2^8 \cdot 2^4 \cdot 2^2 \cdot 2^1$, or 2^{63}. ■

Example 2.6.5 When they were first introduced, postal zip codes were five-digit numbers, theoretically ranging from 00000 to 99999. (In reality, the lowest zip code was 00601 for San Juan, Puerto Rico; the highest was 99950 for Ketchikan, Alaska.) An additional four digits have since been added, so each zip code is now a nine-digit number: an initial five digits, followed by a hyphen, followed by a final four digits.

Let $N(A)$ denote the number of zip codes in the set A, which contains all possible zip codes having at least one 7 among the first five digits; let $N(B)$ denote the number of zip codes in the set B, which contains all possible zip codes having at least one 7 among the final four digits; and let $N(T)$ denote the number of all possible nine-digit zip codes.

Find $N(T)$, $N(A)$, $N(B)$, $N(A \cap B)$, and $N(A \cup B)$. Assume that any digit from 0 to 9 can appear any number of times in a zip code.

Since each of the nine positions in a zip code can be occupied by any of ten digits, $N(T) = 10^9$. Figure 2.6.5 shows examples of zip codes belonging to A, B, $A \cap B$, and $A \cup B$.

$$\begin{array}{ccccccccccccc}
\underline{3} & \underline{7} & \underline{2} & \underline{1} & \underline{7} & - & \underline{4} & \underline{4} & \underline{1} & \underline{6} & \in & A \\
\underline{1} & \underline{6} & \underline{7} & \underline{9} & \underline{4} & - & \underline{0} & \underline{7} & \underline{2} & \underline{1} & \in & B \\
\underline{7} & \underline{0} & \underline{6} & \underline{2} & \underline{1} & - & \underline{7} & \underline{7} & \underline{3} & \underline{7} & \in & A \cap B \\
\underline{2} & \underline{9} & \underline{7} & \underline{5} & \underline{5} & - & \underline{6} & \underline{6} & \underline{7} & \underline{4} & \in & A \cup B
\end{array}$$

Figure 2.6.5

Notice that the number of zip codes in the set A is necessarily equal to the total number of zip codes minus all the zip codes that have *no* 7's in the first five positions. That is,

$$N(A) = N(T) - 9^5 \cdot 10^4 = 10^9 - 9^5 \cdot 10^4$$

Likewise,

$$N(B) = N(T) - 10^5 \cdot 9^4 = 10^9 - 10^5 \cdot 9^4$$

By definition, $A \cap B$ is the set of outcomes having at least one 7 in the first five positions *and* at least one 7 in the last four positions (see Figure 2.6.6). By the

$$\begin{array}{lcc|c}
 & \overbrace{\bar{1}\quad \bar{2}\quad \bar{3}\quad \bar{4}\quad \bar{5}}^{\text{at least one 7}} & & \overbrace{\bar{6}\quad \bar{7}\quad \bar{8}\quad \bar{9}}^{\text{at least one 7}} \\
\text{No. of ways:} & 10^5 - 9^5 & & 10^4 - 9^4
\end{array}$$

Figure 2.6.6

Multiplication rule, then,

$$N(A \cap B) = (10^5 - 9^5) \cdot (10^4 - 9^4)$$

Finally,

$$\begin{aligned} N(A \cup B) &= N(A) + N(B) - N(A \cap B) \\ &= 10^9 - 9^5 \cdot 10^4 + 10^9 - 10^5 \cdot 9^4 - (10^5 - 9^5)(10^4 - 9^4) \\ &= 612{,}579{,}511 \end{aligned}$$

As a partial check, $N(A \cup B)$ should be equal to the total number of zip codes minus the number of zip codes with no 7's. But $10^9 - 9^9 = 612{,}579{,}511$. ■

Problem-Solving Hints

(Doing combinatorial problems)

Combinatorial questions sometimes call for problem-solving techniques that are not routinely used in other areas of mathematics. The three listed below are especially helpful.

1. Draw a diagram that shows the structure of the outcomes that are being counted. Be sure to include (or indicate) all relevant variations. A case in point is Figure 2.6.3. Almost invariably, diagrams such as these will suggest the formula, or combination of formulas, that should be applied.
2. Use enumerations to "test" the appropriateness of a formula. Typically, the answer to a combinatorial problem—that is, the number of ways to do something—will be so large that listing all possible outcomes is not feasible. It often *is* feasible, though, to construct a simple, but analogous, problem for which the entire set of outcomes can be identified (and counted). If the proposed formula does not agree with the simple-case enumeration, we know that our analysis of the original question is incorrect.
3. If the outcomes to be counted fall into structurally different categories, the total number of outcomes will be the *sum* (not the product) of the number of outcomes in each category.

Questions

2.6.1. A chemical engineer wishes to observe the effects of temperature, pressure, and catalyst concentration on the yield resulting from a certain reaction. If she intends to include two different temperatures, three pressures, and two levels of catalyst, how many different runs must she make in order to observe each temperature-pressure-catalyst combination exactly twice?

2.6.2. A coded message from a CIA operative to his Russian KGB counterpart is to be sent in the form Q4ET, where the first and last entries must be consonants; the second, an integer 1 through 9; and the third, one of the six vowels. How many different ciphers can be transmitted?

2.6.3. How many terms will be included in the expansion of

$$(a + b + c)(d + e + f)(x + y + u + v + w)$$

Which of the following will be included in that number: *aeu, cdx, bef, xvw*?

2.6.4. Suppose that the format for license plates in a certain state is two letters followed by four numbers.

(a) How many different plates can be made?

(b) How many different plates are there if the letters can be repeated but no two numbers can be the same?

(c) How many different plates can be made if repetitions of numbers and letters are allowed except that no plate can have four zeros?

2.6.5. How many integers between 100 and 999 have distinct digits, and how many of those are odd numbers?

2.6.6. A fast-food restaurant offers customers a choice of eight toppings that can be added to a hamburger. How many different hamburgers can be ordered?

2.6.7. In baseball there are twenty-four different "base-out" configurations (runner on first—two outs, bases loaded—none out, and so on). Suppose that a new game, sleazeball, is played where there are seven bases (excluding home plate) and each team gets five outs an inning. How many base-out configurations would be possible in sleazeball?

2.6.8. Recall the postal zip codes described in Example 2.6.5.

(a) If viewed as nine-digit numbers, how many zip codes are greater than 700,000,000?

(b) How many zip codes will the digits in the nine positions alternate between even and odd?

(c) How many zip codes will have the first five digits be all different odd numbers and the last four digits be two 2's and two 4's?

2.6.9. A restaurant offers a choice of four appetizers, fourteen entrees, six desserts, and five beverages. How many different meals are possible if a diner intends to order only three courses? (Consider the beverage to be a "course.")

2.6.10. An octave contains twelve distinct notes (on a piano, five black keys and seven white keys). How many different eight-note melodies within a single octave can be written if the black keys and white keys need to alternate?

2.6.11. Residents of a condominium have an automatic garage door opener that has a row of eight buttons. Each garage door has been programmed to respond to a particular set of buttons being pushed. If the condominium houses 250 families, can residents be assured that no two garage doors will open on the same signal? If so, how many additional families can be added before the eight-button code becomes inadequate? (*Note:* The order in which the buttons are pushed is irrelevant.)

2.6.12. In international Morse code, each letter in the alphabet is symbolized by a series of dots and dashes: the letter *a*, for example, is encoded as "· –". What is the minimum number of dots and/or dashes needed to represent any letter in the English alphabet?

2.6.13. The decimal number corresponding to a sequence of n binary digits $a_0, a_1, \ldots, a_{n-1}$, where each a_i is either 0 or 1, is defined to be

$$a_0 2^0 + a_1 2^1 + \cdots + a_{n-1} 2^{n-1}$$

For example, the sequence 0 1 1 0 is equal to 6 $(= 0 \cdot 2^0 + 1 \cdot 2^1 + 1 \cdot 2^2 + 0 \cdot 2^3)$. Suppose a fair coin is tossed nine times. Replace the resulting sequence of H's and T's with a binary sequence of 1's and 0's (1 for H, 0 for T). For how many sequences of tosses will the decimal corresponding to the observed set of heads and tails exceed 256?

2.6.14. Given the letters in the word

$$Z\ O\ M\ B\ I\ E\ S$$

in how many ways can two of the letters be arranged such that one is a vowel and one is a consonant?

2.6.15. Suppose that two cards are drawn—in order—from a standard 52-card poker deck. In how many ways can the first card be a club and the second card be an ace?

2.6.16. Monica's vacation plans require that she fly from Nashville to Chicago to Seattle to Anchorage. According to her travel agent, there are three available flights from Nashville to Chicago, five from Chicago to Seattle, and two from Seattle to Anchorage. Assume that the numbers of options she has for return flights are the same. How many round-trip itineraries can she schedule?

COUNTING PERMUTATIONS (WHEN THE OBJECTS ARE ALL DISTINCT)

Ordered sequences arise in two fundamentally different ways. The first is the scenario addressed by the multiplication rule—a process is comprised of k operations, each allowing n_i options, $i = 1, 2, \ldots, k$; choosing one version of each operation leads to $n_1 n_2 \ldots n_k$ possibilities.

The second occurs when an ordered arrangement of some specified length k is formed from a finite collection of objects. Any such arrangement is referred to as a *permutation of length k*. For example, given the three objects A, B, and C, there are six different permutations of length two that can be formed if the objects cannot be repeated: AB, AC, BC, BA, CA, and CB.

Theorem 2.6.1 *The number of permutations of length k that can be formed from a set of n distinct elements, repetitions not allowed, is denoted by the symbol ${}_nP_k$, where*

$$ {}_nP_k = n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}$$

Proof Any of the n objects may occupy the first position in the arrangement, any of $n-1$ the second, and so on—the number of choices available for filling the kth position will be $n-k+1$ (see Figure 2.6.7). The theorem follows, then, from the multiplication rule: There will be $n(n-1)\cdots(n-k+1)$ ordered arrangements.

Figure 2.6.7

Choices:	n	$n-1$	$\cdots$	$n-(k-2)$	$n-(k-1)$
	1	2		$k-1$	k

Position in sequence

Corollary 2.6.2 *The number of ways to permute an entire set of n distinct objects is* ${}_nP_n = n(n-1)(n-2)\cdots 1 = n!$.

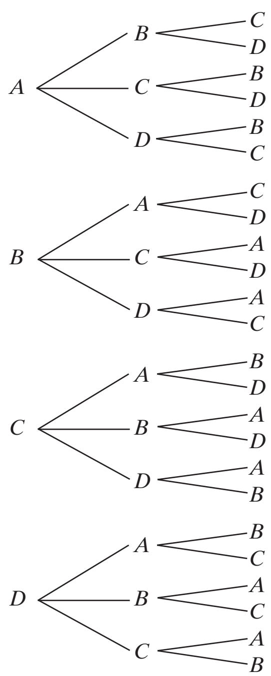
Example 2.6.6 How many permutations of length $k=3$ can be formed from the set of $n=4$ distinct elements, A, B, C, and D?

According to Theorem 2.6.1, the number should be 24:

$$\frac{n!}{(n-k)!} = \frac{4!}{(4-3)!} = \frac{4\cdot 3\cdot 2\cdot 1}{1} = 24$$

Confirming that figure, Table 2.6.2 lists the entire set of twenty-four permutations and illustrates the argument used in the proof of the theorem.

Table 2.6.2



1.	(ABC)
2.	(ABD)
3.	(ACB)
4.	(ACD)
5.	(ADB)
6.	(ADC)
7.	(BAC)
8.	(BAD)
9.	(BCA)
10.	(BCD)
11.	(BDA)
12.	(BDC)
13.	(CAB)
14.	(CAD)
15.	(CBA)
16.	(CBD)
17.	(CDA)
18.	(CDB)
19.	(DAB)
20.	(DAC)
21.	(DBA)
22.	(DBC)
23.	(DCA)
24.	(DCB)

■

Example 2.6.7

In her sonnet with the famous first line, "How do I love thee? Let me count the ways," Elizabeth Barrett Browning listed eight. Suppose Ms. Browning had decided that writing greeting cards afforded her a better format for expressing her feelings. For how many years could she have corresponded with her favorite beau on a daily basis and never sent the same card twice? Assume that each card contains exactly four of the eight "ways" and that order matters.

In selecting the verse for a card, Ms. Browning would be creating a permutation of length $k = 4$ from a set of $n = 8$ distinct objects. According to Theorem 2.6.1,

$$\begin{aligned}\text{number of different cards} = {}_8P_4 = \frac{8!}{(8-4)!} &= 8 \cdot 7 \cdot 6 \cdot 5 \\ &= 1680\end{aligned}$$

At the rate of a card a day, she could have kept the correspondence going for more than four and one-half years. ■

Example 2.6.8

Years ago—long before Rubik's Cubes and electronic games had become epidemic—puzzles were much simpler. One of the more popular combinatorial-related diversions was a four-by-four grid consisting of fifteen movable squares and one empty space. The object was to maneuver as quickly as possible an arbitrary configuration (Figure 2.6.8a) into a specific pattern (Figure 2.6.8b). How many different ways could the puzzle be arranged?

Take the empty space to be square number 16 and imagine the four rows of the grid laid end to end to make a sixteen-digit sequence. Each permutation of that sequence corresponds to a different pattern for the grid. By the corollary to Theorem 2.6.1, the number of ways to position the tiles is 16!, or more than twenty trillion (20,922,789,888,000, to be exact). (*Note:* Not all of the 16! permutations can be generated without physically removing some of the tiles. Think of the two-by-two version of Figure 2.6.8 with tiles numbered 1 through 3. How many of the 4! theoretical configurations can actually be formed?)

13	1	8	7
6	9	3	11
2	10		4
5	12	15	14

(a)

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	

(b)

Figure 2.6.8 ■

Example 2.6.9

A deck of fifty-two cards is shuffled and dealt face up in a row. For how many arrangements will the four aces be adjacent?

This is a good example illustrating the problem-solving benefits that come from drawing diagrams, as mentioned earlier. Figure 2.6.9 shows the basic structure that needs to be considered: The four aces are positioned as a "clump" somewhere between or around the forty-eight non-aces.

Clearly, there are forty-nine "spaces" that could be occupied by the four aces (in front of the first non-ace, between the first and second non-aces, and so on). Furthermore, by the corollary to Theorem 2.6.1, once the four aces are assigned to one of those forty-nine positions, they can still be permuted in ${}_4P_4 = 4!$ ways. Similarly, the

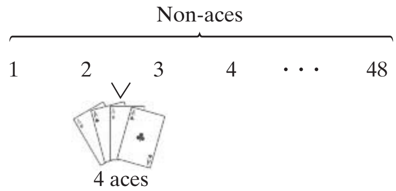


Figure 2.6.9

forty-eight non-aces can be arranged in ${}_{48}P_{48} = 48!$ ways. It follows from the multiplication rule, then, that the number of arrangements having consecutive aces is the product $49 \cdot 4! \cdot 48!$, or, approximately, 1.46×10^{64}. ■

Comment Computing $n!$ can be quite cumbersome, even for n's that are fairly small: We saw in Example 2.6.8, for instance, that 16! is already in the trillions. Fortunately, an easy-to-use approximation is available. According to *Stirling's formula*,

$$n! \doteq \sqrt{2\pi} n^{n+1/2} e^{-n}$$

In practice, we apply Stirling's formula by writing

$$\log_{10}(n!) \doteq \log_{10}\left(\sqrt{2\pi}\right) + \left(n + \frac{1}{2}\right)\log_{10}(n) - n\log_{10}(e)$$

and then exponentiating the right-hand side.

In Example 2.6.9, the number of arrangements was calculated to be $49 \cdot 4! \cdot 48!$, or $24 \cdot 49!$. Substituting into Stirling's formula, we can write

$$\begin{aligned}\log_{10}(49!) &\doteq \log_{10}\left(\sqrt{2\pi}\right) + \left(49 + \frac{1}{2}\right)\log_{10}(49) - 49\log_{10}(e) \\ &\approx 62.783366\end{aligned}$$

Therefore,

$$\begin{aligned}24 \cdot 49! &\doteq 24 \cdot 10^{62.78337} \\ &= 1.46 \times 10^{64}\end{aligned}$$

Example 2.6.10 In chess a rook can move vertically and horizontally (see Figure 2.6.10). It can capture any unobstructed piece located anywhere in its own row or column. In how many ways can eight distinct rooks be placed on a chessboard (having eight rows and eight columns) so that no two can capture one another?

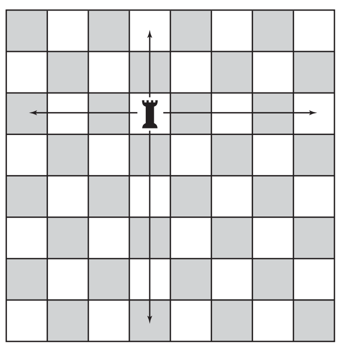


Figure 2.6.10 ■

To start with a simpler problem, suppose that the eight rooks are all identical. Since no two rooks can be in the same row or same column (why?), it follows that each row must contain exactly one. The rook in the first row, however, can be in any of eight columns; the rook in the second row is then limited to being in one of seven columns, and so on. By the multiplication rule, then, the number of noncapturing configurations for eight identical rooks is ${}_8P_8$, or 8! (see Figure 2.6.11).

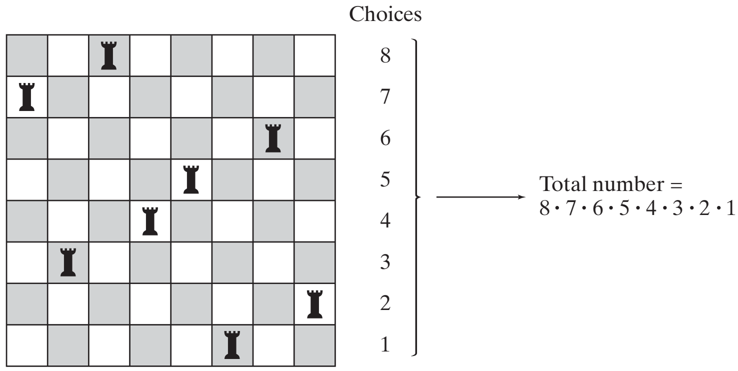


Figure 2.6.11

Now imagine the eight rooks to be distinct—they might be numbered, for example, 1 through 8. The rook in the first row could be marked with any of eight numbers; the rook in the second row with any of the remaining seven numbers; and so on. Altogether, there would be 8! numbering patterns *for each configuration*. The total number of ways to position eight distinct, noncapturing rooks, then, is $8! \cdot 8!$, or 1,625,702,400.

Example 2.6.11

A group of n families, each having m members, are to be lined up in a row for a photograph at a family reunion. In how many ways can those nm individuals be arranged if members of a family must stay together?

Figure 2.6.12 shows one such arrangement. Notice, first of all, that the families—as n distinct groups—can be positioned in $n!$ ways. Also, the individual members

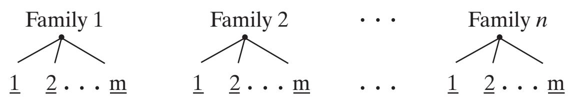


Figure 2.6.12

of *each* family can be rearranged in $m!$ ways. By the multiplication rule, then, the total number of different pictures the photographer could take would be $n!(m!)^n$. Is the latter a big number? Yes, indeed. If only $n = 5$ families attended, each having only $m = 4$ members, the photographer would have almost a *billion* different arrangements to choose from:

$$5!\,(4!)^5 = 955{,}514{,}880$$

Even more amazing, though, is how the number of different arrangements would increase if members of different families were allowed to intermingle. Under those rules, each possible photograph would simply be a permutation of nm people—where in this case $nm = 5(4) = 20$. But the number of permutations of 20 distinct objects is 20!, or 2.4329×10^{18}.

Currently, the total *world* population is approximately eight billion people, or 8.0×10^9. Since

$$2.4329 \times 10^{18}/8.0 \times 10^9 = 3.04 \times 10^8$$

it would be cheaper—in terms of film costs—to take 300,000,000 pictures *of every person living on the planet* than it would be to take the picture of all possible ways for twenty people to be lined up in a row. ■

Example 2.6.12 Consider the set of nine-digit numbers that can be formed by rearranging without repetition the integers 1 through 9. For how many of those permutations will the 1 and the 2 precede the 3 and the 4? That is, we want to count sequences like 7 2 5 1 3 6 9 4 8 but not like 6 8 1 5 4 2 7 3 9.

At first glance, this seems to be a problem well beyond the scope of Theorem 2.6.1. With the help of a symmetry argument, though, its solution is surprisingly simple.

Think of just the digits 1 through 4. By the corollary on p. 72, those four numbers give rise to $4!(= 24)$ permutations. Of those twenty-four, only four—(1, 2, 3, 4), (2, 1, 3, 4), (1, 2, 4, 3), and (2, 1, 4, 3)—have the property that the 1 and the 2 come before the 3 and the 4. It follows that $\frac{4}{24}$ of the total number of nine-digit permutations should satisfy the condition being imposed on 1, 2, 3, and 4. Therefore,

$$\begin{aligned}\text{number of permutations where 1 and 2 precede 3 and 4} &= \frac{4}{24}\cdot 9! \\ &= 60{,}480\end{aligned}$$ ■

Questions

2.6.17. The board of a large corporation has six members willing to be nominated for office. How many different "president/vice president/treasurer" slates could be submitted to the stockholders?

2.6.18. How many ways can a set of four tires be put on a car if all the tires are interchangeable? How many ways are possible if two of the four are snow tires?

2.6.19. Use Stirling's formula to approximate 30!. (*Note:* The exact answer is 265,252,859,812,191,058,636, 308,480,000,000.)

2.6.20. The nine members of the music faculty baseball team, the Mahler Maulers, are all incompetent, and each can play any position equally poorly. In how many different ways can the Maulers take the field?

2.6.21. A three-digit number is to be formed from the digits 1 through 7, with no digit being used more than once. How many such numbers would be less than 289?

2.6.22. Four men and four women are to be seated in a row of chairs numbered 1 through 8.

(a) How many total arrangements are possible?

(b) How many arrangements are possible if the men are required to sit in alternate chairs?

2.6.23. An engineer needs to take three technical electives sometime during his final four semesters. The three are to be selected from a list of ten. In how many ways can he schedule those classes, assuming that he never wants to take more than one technical elective in any given term?

2.6.24. How many ways can a twelve-member cheerleading squad (six men and six women) pair up to form six male-female teams? How many ways can six male-female teams be positioned along a sideline? What might the number $6!6!2^6$ represent? What might the number $6!6!2^6 2^{12}$ represent?

2.6.25. Suppose that a seemingly interminable German opera is recorded on all six sides of a three-record album. In how many ways can the six sides be played so that at least one is out of order?

2.6.26. A new horror movie, *Friday the 13th, Part X*, will star Jason's great-grandson (also named Jason) as a psychotic trying to dispatch (as gruesomely as possible) eight camp counselors, four men and four women. (a) How many scenarios (i.e., victim orders) can the screenwriters devise, assuming they want Jason to do away with all the men before going after any of the women? (b) How many scripts are possible if the only restriction imposed on Jason is that he save Muffy for last?

2.6.27. Suppose that ten people, including you and a friend, line up for a group picture. How many ways can the photographer rearrange the line if she wants to keep exactly three people between you and your friend?

2.6.28. Use an induction argument to prove Theorem 2.6.1. (*Note:* This was the first mathematical result known to have been proved by induction. It was done in 1321 by Levi ben Gerson.)

2.6.29. In how many ways can a pack of fifty-two cards be dealt to thirteen players, four to each, so that every player has one card of each suit?

2.6.30. If the definition of $n!$ is to hold for all nonnegative integers n, show that it follows that $0!$ must equal 1.

2.6.31. The crew of Apollo 17 consisted of a pilot, a copilot, and a geologist. Suppose that NASA had actually trained nine aviators and four geologists as candidates for the flight. How many different crews could they have assembled?

2.6.32. Uncle Harry and Aunt Minnie will both be attending your next family reunion. Unfortunately, they hate each other. Unless they are seated with at least two people between them, they are likely to get into a shouting match. The side of the table at which they will be seated has seven chairs. How many seating arrangements are available for those seven people if a safe distance is to be maintained between your aunt and your uncle?

2.6.33. In how many ways can the digits 1 through 9 be arranged such that

(a) all the even digits precede all the odd digits?

(b) all the even digits are adjacent to each other?

(c) two even digits begin the sequence and two even digits end the sequence?

(d) the even digits appear in either ascending or descending order?

COUNTING PERMUTATIONS (WHEN THE OBJECTS ARE NOT ALL DISTINCT)

The corollary to Theorem 2.6.1 gives a formula for the number of ways an entire set of n objects can be permuted *if the objects are all distinct*. Fewer than $n!$ permutations are possible, though, if some of the objects are identical. For example, there are $3! = 6$ ways to permute the three distinct objects A, B, and C:

ABC
ACB
BAC
BCA
CAB
CBA

If the three objects to permute, though, are A, A, and B—that is, if two of the three are identical—the number of permutations decreases to three:

AAB
ABA
BAA

As we will see, there are many real-world applications where the n objects to be permuted belong to r different categories, each category containing one or more identical objects.

Theorem 2.6.2 *The number of ways to arrange n objects, n_1 being of one kind, n_2 of a second kind, $\ldots$, and n_r of an rth kind, is*

$$\frac{n!}{n_1!\,n_2!\cdots n_r!}$$

where $\sum_{i=1}^{r} n_i = n$.

Proof Let N denote the total number of such arrangements. For any one of those N, the similar objects (if they were actually different) could be arranged in $n_1!\,n_2!\cdots n_r!$ ways. (Why?) It follows that $N \cdot n_1!\,n_2!\cdots n_r!$ is the total number of ways to arrange n (distinct) objects. But $n!$ equals that same number. Setting $N \cdot n_1!\,n_2!\cdots n_r!$ equal to $n!$ gives the result.

Comment Ratios like $n!/(n_1!\,n_2!\cdots n_r!)$ are called *multinomial coefficients* because the general term in the expansion of

$$(x_1 + x_2 + \cdots + x_r)^n$$

is

$$\frac{n!}{n_1!n_2!\cdots n_r!}x_1^{n_1}x_2^{n_2}\cdots x_r^{n_r}$$

Example 2.6.13 A pastry in a vending machine costs 85¢. In how many ways can a customer put in two quarters, three dimes, and one nickel?

Figure 2.6.13

Order in which coins are deposited

If all coins of a given value are considered identical, then a typical deposit sequence, say, $QDDQND$ (see Figure 2.6.13), can be thought of as a permutation of $n = 6$ objects belonging to $r = 3$ categories, where

$$n_1 = \text{number of nickels} = 1$$

$$n_2 = \text{number of dimes} = 3$$

$$n_3 = \text{number of quarters} = 2$$

By Theorem 2.6.2, there are sixty such sequences:

$$\frac{n!}{n_1!n_2!n_3!} = \frac{6!}{1!3!2!} = 60$$

Of course, had we assumed the coins were distinct (having been minted at different places and different times), the number of distinct permutations would have been 6!, or 720. ■

Example 2.6.14 Prior to the seventeenth century there were no scientific journals, a state of affairs that made it difficult for researchers to document discoveries. If a scientist sent a copy of his work to a colleague, there was always a risk that the colleague might claim it as his own. The obvious alternative—wait to get enough material to publish a book—invariably resulted in lengthy delays. So, as a sort of interim documentation, scientists would sometimes send each other anagrams—letter puzzles that, when properly unscrambled, summarized in a sentence or two what had been discovered.

When Christiaan Huygens (1629–1695) looked through his telescope and saw the ring around Saturn, he composed the following anagram (203):

$$aaaaaaa, ccccc, d, eeeee, g, h, iiiiiii, llll, mm,$$

$$nnnnnnnnn, oooo, pp, q, rr, s, ttttt, uuuuu$$

How many ways can the sixty-two letters in Huygens's anagram be arranged?

Let $n_1(= 7)$ denote the number of a's, $n_2(= 5)$ the number of c's, and so on. Substituting into the appropriate multinomial coefficient, we find

$$N = \frac{62!}{7!5!1!5!1!1!7!4!2!9!4!2!1!2!1!5!5!}$$

as the total number of arrangements. To get a feeling for the magnitude of N, we need to apply Stirling's formula to the numerator. Since

$$62! \doteq \sqrt{2\pi}e^{-62}62^{62.5}$$

then

$$\begin{aligned}\log(62!) &\doteq \log\left(\sqrt{2\pi}\right) - 62 \cdot \log(e) + 62.5 \cdot \log(62) \\ &\doteq 85.49731\end{aligned}$$

The antilog of 85.49731 is 3.143×10^{85}, so

$$N \doteq \frac{3.143 \times 10^{85}}{7!5!1!5!1!1!7!4!2!9!4!2!1!2!1!5!5!}$$

is a number on the order of 3.6×10^{60}. Huygens was clearly taking no chances! (*Note:* When appropriately rearranged, the anagram becomes "Annulo cingitur tenui, plano, nusquam cohaerente, ad eclipticam inclinato," which translates to "Surrounded by a thin ring, flat, suspended nowhere, inclined to the ecliptic.") ■

Example 2.6.15

What is the coefficient of x^{23} in the expansion of $(1 + x^5 + x^9)^{100}$?

To understand how this question relates to permutations, consider the simpler problem of expanding $(a + b)^2$:

$$\begin{aligned}(a+b)^2 &= (a+b)(a+b) \\ &= a \cdot a + a \cdot b + b \cdot a + b \cdot b \\ &= a^2 + 2ab + b^2\end{aligned}$$

Notice that each term in the first $(a + b)$ is multiplied by each term in the second $(a+b)$. Moreover, the coefficient that appears in front of each term in the expansion corresponds to the number of ways that term can be formed. For example, the 2 in the term $2ab$ reflects the fact that the product ab can result from two different multiplications:

$$\underbrace{(a+b)(a+b)}_{ab} \qquad \text{or} \qquad (a+\underbrace{b)\,(a}_{ab}+b)$$

By analogy, the coefficient of x^{23} in the expansion of $(1 + x^5 + x^9)^{100}$ will be the number of ways that one term from each of the one hundred factors $(1+x^5+x^9)$ can be multiplied together to form x^{23}. The only factors that will produce x^{23}, though, are the set of two x^9's, one x^5, and ninety-seven 1's:

$$x^{23} = x^9 \cdot x^9 \cdot x^5 \cdot 1 \cdot 1 \cdots 1$$

It follows that the *coefficient* of x^{23} is the number of ways to permute two x^9's, one x^5, and ninety-seven 1's. So, from Theorem 2.6.2,

$$\begin{aligned}\text{coefficient of } x^{23} &= \frac{100!}{2!1!97!} \\ &= 485{,}100\end{aligned}$$

■

Example 2.6.16

A palindrome is a phrase whose letters are in the same order whether they are read backward or forward, such as Napoleon's lament

Able was I ere I saw Elba.

or the often-cited

Madam, I'm Adam.

Words themselves can become the units in a palindrome, as in the sentence

Girl, bathing on Bikini, eyeing boy,

finds boy eyeing bikini on bathing girl.

Suppose the members of a set consisting of four objects of one type, six of a second type, and two of a third type are to be lined up in a row. How many of those permutations are palindromes?

Think of the twelve objects to arrange as being four A's, six B's, and two C's. If the arrangement is to be a palindrome, then half of the A's, half of the B's, and half of the C's must occupy the first six positions in the permutation. Moreover, the final six members of the sequence must be in the reverse order of the first six. For example, if the objects comprising the first half of the permutation were

$$C \quad A \quad B \quad A \quad B \quad B$$

then the last six would need to be in the order

$$B \quad B \quad A \quad B \quad A \quad C$$

It follows that the number of palindromes is the number of ways to permute the first six objects in the sequence, because once the first six are positioned, there is only one arrangement of the last six that will complete the palindrome. By Theorem 2.6.2, then,

$$\text{number of palindromes} = 6!/(2!3!1!) = 60 \qquad \blacksquare$$

Example 2.6.17 A deliveryman is currently at Point X and needs to stop at Point 0 before driving through to Point Y (see Figure 2.6.14). How many different routes can he take without ever going out of his way?

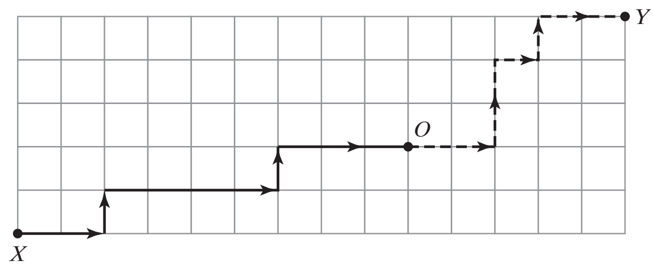


Figure 2.6.14

Notice that any admissible path from, say, X to 0 is an ordered sequence of 11 "moves"—nine east and two north. Pictured in Figure 2.6.14, for example, is the particular X to 0 route

$$E \quad E \quad N \quad E \quad E \quad E \quad E \quad N \quad E \quad E \quad E$$

Similarly, any acceptable path from 0 to Y will necessarily consist of five moves east and three moves north (the one indicated is $E\ E\ N\ N\ E\ N\ E\ E$).

Since each path from X to 0 corresponds to a unique permutation of nine E's and two N's, the *number* of such paths (from Theorem 2.6.2) is the quotient

$$11!/(9!2!) = 55$$

For the same reasons, the number of different paths from 0 to Y is

$$8!/(5!3!) = 56$$

By the multiplication rule, then, the total number of admissible routes from X to Y that pass through 0 is the product of 55 and 56, or 3080. ■

Example 2.6.18

Stacy is worried that her recently-released-from-prison stalker-out-for-revenge ex-boyfriend Psycho Bob is trying to hack into her e-mail account. Crazy, but not stupid, Bob has programmed an algorithm capable of checking three billion passwords a second, which he intends to run 24/7. Suppose Stacy intends to change her password every month. What is the probability her e-mail gets broken into at least once in the next six months?

Stacy's e-mail provider requires 10-digit passwords—four letters (each having two options, lower case or upper case), four numbers, and two symbols (chosen from a list of eight). Figure 2.6.15 shows one such admissible sequence.

$$\underline{7}\quad \underline{3}\quad \underline{\mathbf{B}}\quad \underline{*}\quad \underline{\mathbf{Q}}\quad \underline{\mathbf{a}}\quad \underline{\#}\quad \underline{6}\quad \underline{\mathbf{a}}\quad \underline{1}$$

Figure 2.6.15

Counting the number of admissible passwords is an exercise in using the multiplication rule. Clearly, the *identities* of the four numbers and two symbols can be chosen in $10 \cdot 10 \cdot 10 \cdot 10$ and $8 \cdot 8$ ways, respectively, while the letters can be chosen in $26 \cdot 26 \cdot 26 \cdot 26 \cdot 2^4$ ways. Moreover, by Theorem 2.6.2 the *positions* for the four numbers, two symbols, and four letters can be assigned in $10!/(4!4!2!)$ ways. The total number of different passwords, then, is the product

$$10^4 \cdot 8^2 \cdot (26)^4 \cdot 2^4 \cdot (10!/(4!4!2!)),\ \text{ or } 1.474 \times 10^{16}$$

The probability that Bob's cyberattack identifies whatever password Stacy is using in a given month is simply equal to the number of passwords his algorithm can check in thirty days divided by the total number of passwords admissible. The former is equal to

$$3{,}000{,}000{,}000\,\text{passwords/sec} \times 60\,\text{sec/min} \times 60\ \text{min/hr} \times 24\,\text{hrs/day}$$
$$\times\ 30\,\text{days/mo.} = 7.776 \times 10^{15}\,\text{passwords/mo}$$

so the probability that Stacy's e-mail privacy is compromised in any given month is the quotient

$$7.776 \times 10^{15}/1.474 \times 10^{16},\ \text{ or } 0.53$$

Of course, Bob's success (or failure) in one month is independent of what happens in any other month, so

$$P(\text{e-mail is hacked into at least once in six months}) =$$
$$1 - P(\text{e-mail is } \textit{not} \text{ hacked into for any of the six months})$$
$$= 1 - (0.47)^6 = 0.989$$ ■

CIRCULAR PERMUTATIONS

Thus far the enumeration results we have seen have dealt with what might be called *linear* permutations—objects being lined up in a row. This is the typical context in which permutation problems arise, but sometimes *nonlinear* arrangements of one sort or another need to be counted. The next theorem gives the basic result associated with *circular* permutations.

Theorem 2.6.3 *There are $(n-1)!$ ways to arrange n distinct objects in a circle.*

Proof Fix any object at the "top" of the circle. The remaining $n-1$ objects can then be permuted in $(n-1)!$ ways. Since any arrangement with a different object at the top can be reproduced by simply rotating one of the original $(n-1)!$ permutations, the statement of the theorem holds.

Example 2.6.19 How many different firing orders are theoretically possible in a six-cylinder engine? (If the cylinders are numbered from 1 to 6, a firing order is a list such as 1, 4, 2, 5, 3, 6 giving the sequence in which fuel is ignited in the six cylinders.)

By a direct application of Theorem 2.6.3, the number of distinct firing orders is $(6-1)!$, or 120. ■

Comment According to legend (84), perhaps the first person for whom the problem of arranging objects in a circle took on life or death significance was Flavius Josephus, an early Jewish scholar and historian. In 66 A.D., Josephus found himself a somewhat reluctant leader of an attempt to overthrow the Roman administration in the town of Judea. But the coup failed, and Josephus and forty of his comrades ended up trapped in a cave, surrounded by an angry Roman army.

Faced with the prospect of imminent capture, others in the group were intent on committing mass suicide, but Josephus's devotion to the cause did not extend quite that far. Still, he did not want to appear cowardly, so he proposed an alternate plan: All forty one would arrange themselves in a circle; then, one by one, the group would go around the circle and kill every seventh remaining person, starting with whoever was seated at the head of the circle. That way, only one of them would have to commit suicide, and the entire action would make more of an impact on the Romans.

To his relief, the group accepted his suggestion and began forming a circle. Josephus, who was reputed to have had some genuine mathematical ability, quickly made his way to the twenty-fifth position. Forty murders later, he was the only person left alive!

Is the story true? Maybe yes, maybe no. Any conclusion would be little more than idle speculation. It *is* known, though, that Josephus was the sole survivor of the siege, and that he surrendered and eventually rose to a position of considerable influence in the Roman government. And, whether true or not, the legend has given rise to some mathematical terminology: Cyclic cancellations of a fixed set of numbers having the property that a specified number is left at the end are referred to as *Josephus permutations*.

Questions

2.6.34. Which state name can generate more permutations, TENNESSEE or FLORIDA?

2.6.35. How many numbers greater than four million can be formed from the digits 2, 3, 4, 4, 5, 5, 5?

2.6.36. An interior decorator is trying to arrange a shelf containing eight books, three with red covers, three with blue covers, and two with brown covers.

(a) Assuming the titles and the sizes of the books are irrelevant, in how many ways can she arrange the eight books?

(b) In how many ways could the books be arranged if they were all considered distinct?

(c) In how many ways could the books be arranged if the red books were considered indistinguishable, but the other five were considered distinct?

2.6.37. Four Nigerians (A, B, C, D), three Chinese (#, *, &), and three Greeks (α, β, γ) are lined up at the box office, waiting to buy tickets for the World's Fair.

(a) How many ways can they position themselves if the Nigerians are to hold the first four places in line; the Chinese, the next three; and the Greeks, the last three?

(b) How many arrangements are possible if members of the same nationality must stay together?

(c) How many different queues can be formed?

(d) Suppose a vacationing Martian strolls by and wants to photograph the ten for her scrapbook. A bit myopic, the Martian is quite capable of discerning the more obvious differences in human anatomy but is unable to distinguish one Nigerian (N) from another, one Chinese (C) from another, or one Greek (G) from another. Instead of perceiving a line to be $B*\beta AD\#\&C\alpha\gamma$, for example, she would see $NCGNNCCNGG$. From the Martian's perspective, in how many different ways can the ten funny-looking Earthlings line themselves up?

2.6.38. How many ways can the letters in the word

$$SLUMGULLION$$

be arranged so that the three L's precede all the other consonants?

2.6.39. A tennis tournament has a field of $2n$ entrants, all of whom need to be scheduled to play in the first round. How many different pairings are possible?

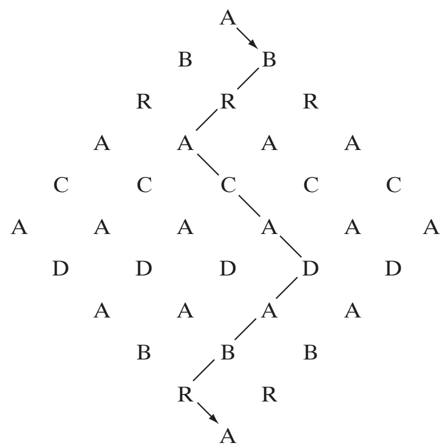
2.6.40. What is the coefficient of x^{12} in the expansion of $(1+x^3+x^6)^{18}$?

2.6.41. In how many ways can the letters of the word

$$ELEEMOSYNARY$$

be arranged so that the S is always immediately followed by a Y?

2.6.42. In how many ways can the word $ABRACADABRA$ be formed in the array pictured above? Assume that the word must begin with the top A and progress diagonally downward to the bottom A.



2.6.43. Suppose a pitcher faces a batter who never swings. For how many different ball/strike sequences will the batter be called out on the fifth pitch?

2.6.44. What is the coefficient of $w^2x^3yz^3$ in the expansion of $(w+x+y+z)^9$?

2.6.45. Imagine six points in a plane, no three of which lie on a straight line. In how many ways can the six points be used as vertices to form two triangles? (*Hint:* Number the points 1 through 6. Call one of the triangles A and the other B. What does the permutation

$$\begin{matrix} A & A & B & B & A & B \\ 1 & 2 & 3 & 4 & 5 & 6 \end{matrix}$$

represent?)

2.6.46. Show that $(k!)!$ is divisible by $k!^{(k-1)!}$. (*Hint:* Think of a related permutation problem whose solution would require Theorem 2.6.2.)

2.6.47. In how many ways can the letters of the word

$$BROBDINGNAGIAN$$

be arranged without changing the order of the vowels?

2.6.48. Make an anagram out of the familiar expression STATISTICS IS FUN. In how many ways can the letters in the anagram be permuted?

2.6.49. Linda is taking a five-course load her first semester: English, math, French, psychology, and history. In how many different ways can she earn three A's and two B's? Enumerate the entire set of possibilities. Use Theorem 2.6.2 to verify your answer.

COUNTING COMBINATIONS

Order is not always a meaningful characteristic of a collection of elements. Consider a poker player being dealt a five-card hand. Whether he receives a 2 of hearts, 4 of clubs, 9 of clubs, jack of hearts, and ace of diamonds *in that order*, or in any one of the other $5!-1$ permutations of those particular five cards is irrelevant, the hand is still the same. As the last set of examples in this section bears out, there are many such situations—problems where our only legitimate concern is with the composition of a set of elements, not with any particular arrangement of them.

We call a collection of k *unordered* elements a *combination of size* k. For example, given a set of $n = 4$ distinct elements—A, B, C, and D—there are *six* ways to form combinations of size 2:

A and B B and C
A and C B and D
A and D C and D

A general formula for counting combinations can be derived quite easily from what we already know about counting permutations.

Theorem 2.6.4 *The number of ways to form combinations of size k from a set of n distinct objects, repetitions not allowed, is denoted by the symbols* $\binom{n}{k}$ *or* ${}_nC_k$*, where*

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

Proof Let the symbol $\binom{n}{k}$ denote the number of combinations satisfying the conditions of the theorem. Since each of those combinations can be ordered in $k!$ ways, the product $k!\binom{n}{k}$ must equal the number of *permutations* of length k that can be formed from n distinct elements. But n distinct elements can be formed into permutations of length k in $n(n-1)\cdots(n-k+1) = n!/(n-k)!$ ways. Therefore,

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

Solving for $\binom{n}{k}$ gives the result.

Comment It often helps to think of combinations in the context of drawing objects out of an urn. If an urn contains n chips labeled 1 through n, the number of ways we can reach in and draw out different samples of size k is $\binom{n}{k}$. In deference to this sampling interpretation for the formation of combinations, $\binom{n}{k}$ is usually read "n things taken k at a time" or "n choose k."

Comment The symbol $\binom{n}{k}$ appears in the statement of a familiar theorem from algebra,

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Since the expression being raised to a power involves two terms, x and y, the constants $\binom{n}{k}$, $k = 0, 1, \ldots, n$, are commonly referred to as *binomial coefficients*.

Example 2.6.20 Eight politicians meet at a fund-raising dinner. How many greetings can be exchanged if each politician shakes hands with every other politician exactly once?

Imagine the politicians to be eight chips—1 through 8—in an urn. A handshake corresponds to an unordered sample of size 2 chosen from that urn. Since repetitions are not allowed (even the most obsequious and overzealous of campaigners would not shake hands with himself!), Theorem 2.6.4 applies, and the total number of handshakes is

$$\binom{8}{2} = \frac{8!}{2!6!}$$

or 28. ■

Example 2.6.21 A chemist is trying to synthesize a part of a straight-chain aliphatic hydrocarbon polymer that consists of twenty-one radicals—ten ethyls (E), six methyls (M), and five propyls (P). Assuming all arrangements of radicals are physically possible, how many different polymers can be formed if no two of the methyl radicals are to be adjacent?

Imagine arranging the E's and the P's without the M's. Figure 2.6.16 shows one such possibility. Consider the sixteen "spaces" between and outside the E's and P's as indicated by the arrows in Figure 2.6.16. In order for the M's to be nonadjacent, they must occupy any six of these locations. But those six spaces can be chosen in $\binom{16}{6}$ ways. And for each of the $\binom{16}{6}$ positionings of the M's, the E's and P's can be permuted in $\frac{15!}{10!5!}$ ways (Theorem 2.6.2).

$$\begin{array}{cccccccccccccccc} & E & E & P & P & E & E & E & P & E & P & E & P & E & E & E \\ \uparrow & \uparrow & \uparrow & \cdots & & & & & & & & & & \cdots & \uparrow & \uparrow \end{array}$$

Figure 2.6.16

So, by the multiplication rule, the total number of polymers having nonadjacent methyl radicals is 24,048,024:

$$\binom{16}{6} \cdot \frac{15!}{10!5!} = \frac{16!}{10!6!} \frac{15!}{10!5!} = (8008)(3003) = 24,048,024$$ ■

Example 2.6.22 Binomial coefficients have many interesting properties. Perhaps the most familiar is Pascal's triangle,[1] a numerical array where each entry is equal to the sum of the two numbers appearing diagonally above it (see Figure 2.6.17). Notice that each entry in Pascal's triangle can be expressed as a binomial coefficient, and the relationship just described appears to reduce to a simple equation involving those coefficients:

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1} \tag{2.6.1}$$

Prove that Equation 2.6.1 holds for all positive integers n and k.

$$\begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1 \\ & \vdots & & & & & & \vdots & \end{array} \qquad \begin{array}{c} \underline{\text{Row}} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ \vdots \end{array} \qquad \begin{array}{ccccccccc} & & & & \binom{0}{0} & & & & \\ & & & \binom{1}{0} & & \binom{1}{1} & & & \\ & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & \\ & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & \\ \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} \\ & \vdots & & & & & & \vdots & \end{array}$$

Figure 2.6.17

[1] Despite its name, Pascal's triangle was not discovered by Pascal. Its basic structure had been known hundreds of years before the French mathematician was born. It was Pascal, though, who first made extensive use of its properties.

Consider a set of $n+1$ distinct objects $A_1, A_2, \ldots, A_{n+1}$. We can obviously draw samples of size k from that set in $\binom{n+1}{k}$ different ways. Now, consider any particular object—for example, A_1. Relative to A_1, each of those $\binom{n+1}{k}$ samples belongs to one of two categories: those containing A_1 and those not containing A_1. To form samples containing A_1, we need to select $k-1$ additional objects from the remaining n. This can be done in $\binom{n}{k-1}$ ways. Similarly, there are $\binom{n}{k}$ ways to form samples not containing A_1. Therefore, $\binom{n+1}{k}$ must equal $\binom{n}{k}+\binom{n}{k-1}$. ■

Example 2.6.23 The answers to combinatorial questions can sometimes be obtained using quite different approaches. What invariably distinguishes one solution from another is the way in which outcomes are characterized.

For example, suppose you have just ordered a roast beef sub at a sandwich shop, and now you need to decide which, if any, of the available toppings (lettuce, tomato, onions, etc.) to add. If the shop has eight "extras" to choose from, how many different subs can you order?

One way to answer this question is to think of each sub as an ordered sequence of length eight, where each position in the sequence corresponds to one of the toppings. At each of those positions, you have two choices—"add" or "do not add" that particular topping. Pictured in Figure 2.6.18 is the sequence corresponding to the sub that has lettuce, tomato, and onion but no other toppings. Since two choices ("add" or "do not add") are available for each of the eight toppings, the multiplication rule tells us that the number of different roast beef subs that could be requested is 2^8, or 256.

Add?							
Y	Y	Y	N	N	N	N	N
Lettuce	Tomato	Onion	Mustard	Relish	Mayo	Pickles	Peppers

Figure 2.6.18

An ordered sequence of length 8, though, is not the only model capable of characterizing a roast beef sandwich. We can also distinguish one roast beef sub from another by the particular *combination* of toppings that each one has. For example, there are $\binom{8}{4} = 70$ different subs having exactly four toppings. It follows that the total number of different sandwiches is the total number of different combinations of size k, where k ranges from 0 to 8. Reassuringly, that sum agrees with the ordered sequence answer:

$$\begin{aligned}\text{total number of different roast beef subs} &= \binom{8}{0}+\binom{8}{1}+\binom{8}{2}+\cdots+\binom{8}{8}\\ &= 1+8+28+\cdots+1\\ &= 256\end{aligned}$$

What we have just illustrated here is another property of binomial coefficients, namely, that

$$\sum_{k=0}^{n}\binom{n}{k} = 2^n \tag{2.6.2}$$

The proof of Equation 2.6.2 is a direct consequence of Newton's binomial expansion (see the second comment following Theorem 2.6.4). ■

Example 2.6.24

Recall Theorem 2.3.7, the formula for the probability of the union of n events, $P(A_1 \cup A_2 \cup \cdots \cup A_n)$. As the special-case discussion on p. 28 indicated, it needs to be proved in general that the formula adds—*once and only once*—the probability of every outcome represented in the Venn diagram for $A_1 \cup A_2 \cup \cdots \cup A_n$.

To that end, consider the set of outcomes in $A_1 \cup A_2 \cup \cdots \cup A_n$ that belong to a specified k of the A_i's *and to no others*. If it can be established that the right-hand-side of Theorem 2.3.7 counts that particular set of outcomes *one* time, the theorem is proved since k was arbitrary.

Consider such a set of outcomes. Relative to the summations making up the right-hand side of Theorem 2.3.7, those outcomes get counted $\binom{k}{1}$ times in $\sum_{i=1}^{n} P(A_i)$, $\binom{k}{2}$ times in $\sum_{i<j} P(A_i \cap A_j)$, $\binom{k}{3}$ times in $\sum_{i<j<k} P(A_i \cap A_j \cap A_k)$, and so on. According to the theorem, then, they will be counted a total of

$$\binom{k}{1} - \binom{k}{2} + \binom{k}{3} - \cdots + (-1)^{k+1}\binom{k}{k}$$

times.

Expressing that sum as a binomial expansion, we can write

$$\begin{aligned}(-1+1)^k = 0^k &= \sum_{j=0}^{k} \binom{k}{j} (-1)^j (1)^{k-j} \\ &= \binom{k}{0} - \binom{k}{1} + \binom{k}{2} - \cdots + (-1)^k \binom{k}{k}\end{aligned}$$

or, equivalently,

$$\binom{k}{1} - \binom{k}{2} + \cdots + (-1)^{k+1}\binom{k}{k} = \binom{k}{0} = 1$$

and the theorem is proved. ■

Example 2.6.25

In Example 2.6.23, the fact that $\sum_{k=0}^{n} \binom{n}{k}$ is equal to 2^n was established by showing that both expressions count the number of ways to complete the same task. Analytically, Equation 2.6.2 could have been derived by simplifying the expansion of $(x + y)^n$. That is,

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Letting $x = y = 1$ gives

$$(1+1)^n = 2^n = \sum_{k=0}^{n} \binom{n}{k} 1^k \cdot 1^{n-k} = \sum_{k=0}^{n} \binom{n}{k}$$

Concluding this section are two other examples of binomial coefficient identities. The first is proven analytically; the second is done by appealing to a sampling experiment.

a. Prove that

$$\binom{n}{1} + 2\binom{n}{2} + \cdots + n\binom{n}{n} = n2^{n-1}$$

Consider the expansion of $(1 + x)^n$:

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k (1)^{n-k} \tag{2.6.3}$$

Differentiating both sides of Equation 2.6.3 with respect to x gives

$$n(1+x)^{n-1} = \sum_{x=0}^{n} \binom{n}{k} kx^{k-1} \tag{2.6.4}$$

Let $x = 1$. Then Equation 2.6.4 reduces to

$$n \cdot 2^{n-1} = 1\binom{n}{1} + 2\binom{n}{2} + \cdots + n\binom{n}{n}$$

b. Prove that

$$\binom{n}{0}^2 + \binom{n}{1}^2 + \cdots + \binom{n}{n}^2 = \binom{2n}{n}$$

Imagine drawing a sample of size n from a population of $2n$ objects, divided into a first set of n objects and a second set of n objects. The desired sample could be formed by drawing zero objects from the first set and n objects from the second set *or* one object from the first set and $n-1$ objects from the second set, and so on. Clearly

$$\binom{n}{0}\binom{n}{n} + \binom{n}{1}\binom{n}{n-1} + \binom{n}{2}\binom{n}{n-2} + \cdots + \binom{n}{n}\binom{n}{0} = \binom{2n}{n}$$

But $\binom{n}{k} = \binom{n}{n-k}$ for all k, so

$$\binom{n}{0}^2 + \binom{n}{1}^2 + \cdots + \binom{n}{n}^2 = \binom{2n}{n}$$

■

Questions

2.6.50. How many straight lines can be drawn between five points (A, B, C, D, and E), no three of which are collinear?

2.6.51. The Alpha Beta Zeta sorority is trying to fill a pledge class of nine new members during fall rush. Among the twenty-five available candidates, fifteen have been judged marginally acceptable and ten highly desirable. How many ways can the pledge class be chosen to give a two-to-one ratio of highly desirable to marginally acceptable candidates?

2.6.52. A boat has a crew of eight: Two of those eight can row only on the stroke side, while three can row only on the bow side. In how many ways can the two sides of the boat be manned?

2.6.53. Nine students, five men and four women, interview for four summer internships sponsored by a city newspaper.

(a) In how many ways can the newspaper choose a set of four interns?

(b) In how many ways can the newspaper choose a set of four interns if it must include two men and two women in each set?

(c) How many sets of four can be picked such that not everyone in a set is of the same sex?

2.6.54. The final exam in History 101 consists of five essay questions that the professor chooses from a pool of seven that are given to the students a week in advance. For how many possible sets of questions does a student need to be prepared? In this situation, does order matter?

2.6.55. Ten basketball players meet in the school gym for a pickup game. How many ways can they form two teams of five each?

2.6.56. Your statistics teacher announces a twenty-page reading assignment on Monday that is to be finished by Thursday morning. You intend to read the first x_1 pages Monday, the next x_2 pages Tuesday, and the final x_3 pages Wednesday, where $x_1 + x_2 + x_3 = 20$, and each $x_i \geq 1$. In how many ways can you complete the assignment? That is, how many different sets of values can be chosen for x_1, x_2, and x_3?

2.6.57. In how many ways can the letters in

$$M\,I\,S\,S\,I\,S\,S\,I\,P\,P\,I$$

be arranged so that no two I's are adjacent?

2.6.58. Prove that

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

directly without appealing to any combinatorial arguments.

2.6.59. Find a recursion formula for $\binom{n}{k+1}$ in terms of $\binom{n}{k}$.

2.6.60. Show that $n(n-1)2^{n-2} = \sum_{k=2}^{n} k(k-1)\binom{n}{k}$.

2.6.61. Prove that successive terms in the sequence $\binom{n}{0}$, $\binom{n}{1}, \ldots, \binom{n}{n}$ first increase and then decrease. [*Hint:* Examine the ratio of two successive terms, $\binom{n}{j+1} \Big/ \binom{n}{j}$.]

2.6.62. Mitch is trying to add a little zing to his cabaret act by telling four jokes at the beginning of each show. His current engagement is booked to run four months. If he gives one performance a night and never wants to repeat the same set of jokes on any two nights, what is the minimum number of jokes he needs in his repertoire?

2.6.63. Compare the coefficients of t^k in $(1+t)^d(1+t)^e = (1+t)^{d+e}$ to prove that

$$\sum_{j=0}^{k} \binom{d}{j}\binom{e}{k-j} = \binom{d+e}{k}$$

2.7 Combinatorial Probability

In Section 2.6 our concern focused on counting the number of ways a given operation, or sequence of operations, could be performed. In Section 2.7 we want to couple those enumeration results with the notion of probability. Putting the two together makes a lot of sense—there are many combinatorial problems where an enumeration, by itself, is not particularly relevant. A poker player, for example, is not interested in knowing the total *number* of ways he can draw to a straight; he *is* interested, though, in his *probability* of drawing to a straight.

In a combinatorial setting, making the transition from an enumeration to a probability is easy. If there are n ways to perform a certain operation and a total of m of those satisfy some stated condition—call it A—then $P(A)$ is defined to be the ratio m/n. This assumes, of course, that all possible outcomes are equally likely.

Historically, the "m over n" idea is what motivated the early work of Pascal, Fermat, and Huygens (recall Section 1.3). Today we recognize that not all probabilities are so easily characterized. Nevertheless, the m/n model—the so-called *classical* definition of probability—is entirely appropriate for describing a wide variety of phenomena.

Example 2.7.1

An urn contains eight chips, numbered 1 through 8. A sample of three is drawn without replacement. What is the probability that the largest chip in the sample is a 5?

Let A be the event "Largest chip in sample is a 5." Figure 2.7.1 shows what must happen in order for A to occur: (1) the 5 chip must be selected, and (2) two chips must be drawn from the subpopulation of chips numbered 1 through 4. By the multiplication rule, the number of samples satisfying event A is the product $\binom{1}{1} \cdot \binom{4}{2}$.

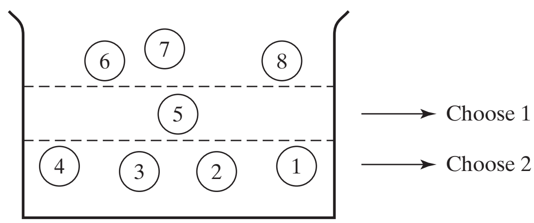


Figure 2.7.1

The sample space S for the experiment of drawing three chips from the urn contains $\binom{8}{3}$ outcomes, all equally likely. In this situation, then, $m = \binom{1}{1} \cdot \binom{4}{2}$, $n = \binom{8}{3}$, and

$$P(A) = \frac{\binom{1}{1} \cdot \binom{4}{2}}{\binom{8}{3}}$$
$$= 0.11$$ ■

Example 2.7.2

An urn contains n red chips numbered 1 through n, n white chips numbered 1 through n, and n blue chips numbered 1 through n (see Figure 2.7.2). Two chips are drawn at random and without replacement. What is the probability that the two drawn are either the same color or the same number?

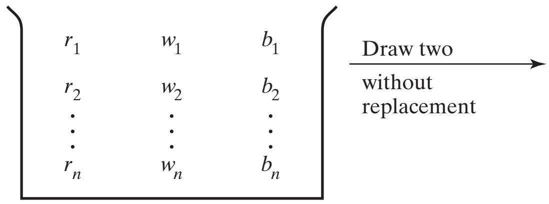


Figure 2.7.2

Let A be the event that the two chips drawn are the same color; let B be the event that they have the same number. We are looking for $P(A \cup B)$.

Since A and B here are mutually exclusive,

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

With $3n$ chips in the urn, the total number of ways to draw an unordered sample of size 2 is $\binom{3n}{2}$. Moreover,

$$\begin{aligned} P(A) &= P(2\text{ reds } \cup 2\text{ whites } \cup 2\text{ blues}) \\ &= P(2\text{ reds}) + P(2\text{ whites}) + P(2\text{ blues}) \\ &= 3\binom{n}{2} \Big/ \binom{3n}{2} \end{aligned}$$

and

$$\begin{aligned} P(B) &= P(\text{two 1's} \cup \text{ two 2's} \cup \cdots \cup \text{two } n\text{'s}) \\ &= n\binom{3}{2} \Big/ \binom{3n}{2} \end{aligned}$$

Therefore,

$$\begin{aligned} P(A \cup B) &= \frac{3\binom{n}{2} + n\binom{3}{2}}{\binom{3n}{2}} \\ &= \frac{n+1}{3n-1} \end{aligned}$$

■

Example 2.7.3

Twelve fair dice are rolled. What is the probability that

a. the first six dice all show one face and the last six dice all show a second face?

b. not all the faces are the same?

c. each face appears exactly twice?

a. The sample space that corresponds to the "experiment" of rolling twelve dice is the set of ordered sequences of length twelve, where the outcome at every position in the sequence is one of the integers 1 through 6. If the dice are fair, all 6^{12} such sequences are equally likely.

Let A be the set of rolls where the first six dice show one face and the second six show another face. Figure 2.7.3 shows one of the sequences in the event A. Clearly, the face that appears for the first half of the sequence could be any of the six integers from 1 through 6.

Faces

2	2	2	2	2	2	4	4	4	4	4	4
1	2	3	4	5	6	7	8	9	10	11	12

Position in sequence

Figure 2.7.3

Five choices would be available for the last half of the sequence (since the two faces cannot be the same). The number of sequences in the event A, then, is ${}_6P_2 = 6 \cdot 5 = 30$. Applying the "m/n" rule gives

$$P(A) = 30/6^{12} = 1.4 \times 10^{-8}$$

b. Let B be the event that not all the faces are the same. Then

$$\begin{aligned} P(B) &= 1 - P(B^C) \\ &= 1 - 6/12^6 \end{aligned}$$

since there are six sequences—(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,), . . ., (6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6,)—where the twelve faces *are* all the same.

c. Let C be the event that each face appears exactly twice. From Theorem 2.6.2, the *number* of ways each face can appear exactly twice is $12!/(2! \cdot 2! \cdot 2! \cdot 2! \cdot 2! \cdot 2!)$. Therefore,

$$\begin{aligned} P(C) &= \frac{12!/(2! \cdot 2! \cdot 2! \cdot 2! \cdot 2! \cdot 2!)}{6^{12}} \\ &= 0.0034 \end{aligned}$$

■

Example 2.7.4 A fair die is tossed n times. What is the probability that the sum of the faces showing is $n + 2$?

The sample space associated with rolling a die n times has 6^n outcomes, all of which in this case are equally likely because the die is presumed fair. There are two "types" of outcomes that will produce a sum of $n + 2$: (a) $n - 1$ 1's and one 3 and (b) $n - 2$ 1's and two 2's (see Figure 2.7.4). By Theorem 2.6.2 the number of sequences having $n - 1$ 1's and one 3 is $\frac{n!}{1!(n-1)!} = n$; likewise, there are $\frac{n!}{2!(n-2)!} = \binom{n}{2}$ outcomes having $n - 2$ 1's and two 2's. Therefore,

$$P(\text{sum} = n + 2) = \frac{n + \binom{n}{2}}{6^n}$$

■

$$\overbrace{\frac{1}{1} \quad \frac{1}{2} \quad \frac{1}{3} \quad \cdots \quad \frac{1}{n-1} \quad \frac{3}{n}}^{\text{Sum} = n+2} \qquad \overbrace{\frac{1}{1} \quad \frac{1}{2} \quad \frac{1}{3} \quad \cdots \quad \frac{1}{n-2} \quad \frac{2}{n-1} \quad \frac{2}{n}}^{\text{Sum} = n+2}$$

Figure 2.7.4

Example 2.7.5 Two monkeys, Mickey and Marian, are strolling along a moonlit beach when Mickey sees an abandoned Scrabble set. Investigating, he notices that some of the letters are missing, and what remain are the following fifty-nine tiles:

A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*
4	1	2	2	7	1	1	3	5	0	3	5	1

N	*O*	*P*	*Q*	*R*	*S*	*T*	*U*	*V*	*W*	*X*	*Y*	*Z*
3	2	0	0	2	8	4	2	0	1	0	2	0

Mickey, being of a romantic bent, would like to impress Marian, so he rearranges the letters hoping to spell something endearing. For some unknown reason, Marian can read, but Mickey is dumb as dirt, so all he can do is scramble the fifty-nine tiles at random and hope for the best. What is the probability he gets lucky and spells out

> She walks in beauty, like the night
> Of cloudless climes and starry skies

As we might imagine, Mickey would have to get *very* lucky. The total number of ways to permute fifty-nine letters—four A's, one B, two C's, and so on—is a direct application of Theorem 2.6.2:

$$\frac{59!}{4!1!2!\ldots 2!0!}$$

But of that number of ways, only one is the couplet he is hoping for. So, since he is arranging the letters randomly, making all permutations equally likely, the probability of his spelling out Byron's lines is

$$\frac{1}{\dfrac{59!}{4!1!2!\ldots 2!0!}}$$

or, using Stirling's formula, about 1.7×10^{-61}. Love may conquer all, but it won't beat those odds: Mickey would be well advised to start working on Plan B. ■

Example 2.7.6 Suppose k people are selected at random. What are the chances that at least two of those k were born on the same day of the year? Known as the *birthday problem*, this is a particularly intriguing example of combinatorial probability because its statement is so simple, its analysis is straightforward, yet its solution is strongly contrary to our intuition.

Picture the k individuals lined up in a row to form an ordered sequence. If leap year is omitted, each person might have any of 365 birthdays. By the multiplication rule, the group as a whole generates a sample space of 365^k birthday sequences (see Figure 2.7.5).

$$\text{Possible birthdays: } \underbrace{\frac{(365)}{1}\quad \frac{(365)}{2}\quad \frac{\cdots}{\cdots}\quad \frac{(365)}{k}}_{\text{Person}} \longrightarrow 365^k \text{ different sequences}$$

Figure 2.7.5

Define A to be the event "At least two people have the same birthday." If each person is assumed to have the same chance of being born on any given day, the 365^k sequences in Figure 2.7.5 are equally likely, and

$$P(A) = \frac{\text{number of sequences in } A}{365^k}$$

Counting the number of sequences in the numerator here is prohibitively difficult because of the complexity of the event A; fortunately, counting the number of sequences in A^c is quite easy. Notice that each birthday sequence in the sample space belongs to exactly one of two categories (see Figure 2.7.6):

1. At least two people have the same birthday.
2. All k people have different birthdays.

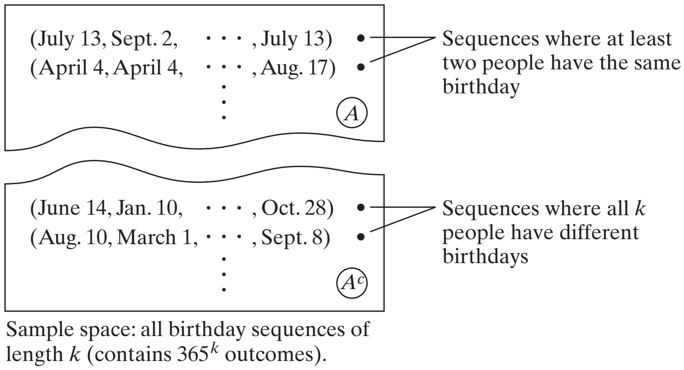


Figure 2.7.6

It follows that

$$\text{number of sequences in } A = 365^k - \text{number of sequences where all } k \text{ people have different birthdays}$$

The number of ways to form birthday sequences for k people subject to the restriction that all k birthdays must be different is simply the number of ways to form permutations of length k from a set of 365 distinct objects:

$$_{365}P_k = 365(364)\cdots(365-k+1)$$

Therefore,

$$\begin{aligned} P(A) &= P(\text{At least two people have the same birthday}) \\ &= \frac{365^k - 365(364)\cdots(365-k+1)}{365^k} \end{aligned}$$

Table 2.7.1 shows $P(A)$ for k values of 15, 22, 23, 40, 50, and 70. Notice how the $P(A)$'s greatly exceed what our intuition would suggest.

Table 2.7.1

k	$P(A) = P$ (At least two have same birthday)
15	0.253
22	0.476
23	0.507
40	0.891
50	0.970
70	0.999

■

Comment The values for $P(A)$ in Table 2.7.1 are actually slight *underestimates* for the true probabilities that at least two of k people will be born on the same day. The assumption made earlier that all 365^k birthday sequences are equally likely is not entirely correct: Births are somewhat more common during the summer than they are during the winter. It has been proven, though, that any sort of deviation from the equally likely model will serve only to *increase* the chances that two or more people will share the same birthday. So, if $k = 40$, for example, the probability is slightly greater than 0.891 that at least two were born on the same day.

Comment Presidential biographies offer one opportunity to "confirm" the unexpectedly large values that Table 2.7.1 gives for $P(A)$. Among our first $k = 40$ presidents, two did have the same birthday: Harding and Polk were both born on November 2. More surprising, though, are the death dates of the presidents: John Adams, Jefferson, and Monroe all died on July 4, and Fillmore and Taft both died on March 8.

Example 2.7.7 One of the more instructive—and to some, one of the more useful—applications of combinatorics is the calculation of probabilities associated with various poker hands. It will be assumed in what follows that five cards are dealt from a poker deck and that no other cards are showing, although some may already have been dealt. The sample space is the set of $\binom{52}{5} = 2{,}598{,}960$ different hands, each having probability $1/2{,}598{,}960$. What are the chances of being dealt (a) a *full house*, (b) *one pair*, and (c) a *straight*? [Probabilities for the various other kinds of poker hands (two pairs, three-of-a-kind, flush, and so on) are gotten in much the same way.]

a. *Full house*. A full house consists of three cards of one denomination and two of another. Figure 2.7.7 shows a full house consisting of three 7's and two queens. Denominations for the three-of-a-kind can be chosen in $\binom{13}{1}$ ways. Then, given that a denomination has been decided on, the three requisite suits can be selected in $\binom{4}{3}$ ways. Applying the same reasoning to the pair gives $\binom{12}{1}$ available denominations, each having $\binom{4}{2}$ possible choices of suits. Thus, by the multiplication rule,

$$P(\text{full house}) = \frac{\binom{13}{1}\binom{4}{3}\binom{12}{1}\binom{4}{2}}{\binom{52}{5}} = 0.00144$$

	2	3	4	5	6	7	8	9	10	J	Q	K	A
D													
H						×					×		
C						×							
S						×					×		

Figure 2.7.7

b. *One pair*. To qualify as a one-pair hand, the five cards must include two of the same denomination and three "single" cards—cards whose denominations match neither the pair nor each other. Figure 2.7.8 shows a pair of 6's. For the pair, there are $\binom{13}{1}$ possible denominations and, once selected, $\binom{4}{2}$ possible suits. Denominations for the three single cards can be chosen $\binom{12}{3}$ ways (see Question 2.7.16), and each card can have any of $\binom{4}{1}$ suits. Multiplying these factors together and dividing by $\binom{52}{2}$ gives a probability of 0.42:

$$P(\text{one pair}) = \frac{\binom{13}{1}\binom{4}{2}\binom{12}{3}\binom{4}{1}\binom{4}{1}\binom{4}{1}}{\binom{52}{5}} = 0.42$$

	2	3	4	5	6	7	8	9	10	J	Q	K	A
D			×										×
H					×		×						
C					×								
S													

Figure 2.7.8

c. *Straight*. A straight is five cards having consecutive denominations *but not all in the same suit*—for example, a 4 of diamonds, 5 of hearts, 6 of hearts, 7 of clubs, and 8 of diamonds (see Figure 2.7.9). An ace may be counted "high" or "low," which means that (10, jack, queen, king, ace) is a straight and so is (ace, 2, 3, 4, 5). (If five consecutive cards are all in the same suit, the hand is called a *straight flush*. The latter is considered a fundamentally different type of hand in the sense that a straight flush "beats" a straight.) To get the numerator for $P(\text{straight})$, we will first ignore the condition that all five cards not be in the same suit and simply count the number of hands having consecutive denominations. Note there are ten sets of consecutive denominations of length five: (ace, 2, 3, 4, 5), (2, 3, 4, 5, 6), . . ., (10, jack, queen, king, ace). With no restrictions on the suits, each card can be either a diamond, heart, club, or spade. It follows, then, that the number of five-card hands having consecutive denominations is $10 \cdot \binom{4}{1}^5$. But forty $(= 10 \cdot 4)$ of those hands are straight flushes. Therefore,

$$P(\text{straight}) = \frac{10 \cdot \binom{4}{1}^5 - 40}{\binom{52}{5}} = 0.00392$$

Table 2.7.2 shows the probabilities associated with all the different poker hands. Hand i beats hand j if $P(\text{hand } i) < P(\text{hand } j)$.

	2	3	4	5	6	7	8	9	10	J	Q	K	A
D			×				×						
H				×	×								
C						×							
S													

Figure 2.7.9

Table 2.7.2

Hand	Probability
One pair	0.42
Two pairs	0.048
Three-of-a-kind	0.021
Straight	0.0039
Flush	0.0020
Full house	0.0014
Four-of-a-kind	0.00024
Straight flush	0.000014
Royal flush	0.0000015

■

Example 2.7.8

A somewhat inebriated conventioneer finds himself in the embarrassing position of being unable to discern whether he is walking forward or backward—or, what is worst, to predict in which of those directions his next step will be. If he is equally likely to walk forward or backward, what is the probability that after hazarding n such maneuvers, he will have moved forward a distance of r steps?

Let x denote the number of steps he takes forward and y denote the number backward. Then

$$x + y = n$$

and

$$x - y = r$$

Solving these equations simultaneously, we get $x = (n+r)/2$ and $y = (n-r)/2$. Thus, out of the 2^n total ways he can take n steps, the number of permutations for which he ends up r steps forward is $n!/\left[\left(\frac{n+r}{2}\right)!\left(\frac{n-r}{2}\right)!\right]$ (recall Theorem 2.6.2). The probability that he shows a net advance of r steps is the quotient

$$\frac{\binom{n}{\frac{n+r}{2}}}{2^n}$$ ■

Comment Example 2.7.8 is describing a one-dimensional *random walk* problem. Over the years, the mathematical properties of random walks in various numbers of dimensions have been a useful tool for modeling such disparate phenomena as Brownian motion, the movements of individual animals and populations of animals, and daily fluctuations in the stock market. The term *random walk* was coined in 1905 by the renowned British statistician, Karl Pearson.

Example 2.7.9

To a purist, buying a LOTTO ticket at your local supermarket is not quite the gambling equivalent of playing a round of Texas Hold'Em in some smoke-filled roadhouse, but there is no denying that the amount of money spent on lotteries each year (estimated to be upwards of \$50 billion in the United States alone) is major-league. And to be fair, lotteries do have an interesting history that goes back thousands of years, long before the first poker game was ever played. Evidence suggests that lotteries helped finance the building of the Great Wall of China. Closer to home, the first permanent English settlement on the North American continent–Jamestown, founded in 1607–was funded by the Virginia Company of London, a group of investors who acquired the rights to sponsor the expedition by winning a lottery.

In recent years, one of the most popular versions of "drawing lots" has been Powerball. For the price of a ticket, a player gets to pick five (distinct) numbers from the integers 1 through 59 and one additional number from the integers 1 through 35. Then at a regularly-scheduled time, five white balls are drawn from a drum containing fifty-nine balls, numbered 1 through 59, and a sixth selection (the Powerball) is drawn from a second drum containing thirty-five balls, numbered 1 through 35. If and how much a player wins depends on the nature of the matches between the ticket numbers picked and the balls actually drawn.

For example, the fourth prize is worth \$100 and awarded to anyone whose ticket correctly matches four of the five white balls but not the Powerball. Calculating the probability of that happening is a straightforward exercise in applying the multiplication rule to the numbers of ways to form various combinations. Since

$$\binom{59}{5} = \text{number of ways to choose five white balls from the first drum}$$

and

$$\binom{35}{1} = \text{number of ways to choose a Powerball from the second drum}$$

the product $\binom{59}{5}\binom{35}{1} = 175,223,510$ is the total number of (equally-likely) outcomes possible in a LOTTO drawing. Your particular ticket has the potential to (1) match exactly four of the white balls drawn in $\binom{5}{4} = 5$ different ways, (2) *not* match the fifth white ball in $\binom{54}{1} = 54$ different ways and (3) *not* match the Powerball in $\binom{34}{1} = 34$ different ways. The probability, then, that your ticket will win the \$100 prize is the ratio

$$\binom{5}{4}\binom{54}{1}\binom{34}{1}\Big/\left[\binom{59}{5}\binom{35}{1}\right] = 9180/175{,}223{,}510 = 0.00005239$$

The entire set of winning combinations, associated probabilities, and payoffs are listed in Table 2.7.3. Adding the entries in the middle column would show that your probability of winning *something* is 0.0314. ■

Table 2.7.3

Winning Pick	Probability	Payoff
1. Match all five white balls *and* Powerball	$[\binom{5}{5}\binom{1}{1}]/[\binom{59}{5}\binom{35}{1}]$	Depends on the number of winners
2. Match all five white balls but not Powerball	$[\binom{5}{5}\binom{34}{1}]/[\binom{59}{5}\binom{35}{1}]$	\$1,000,000
3. Match four white balls *and* Powerball	$[\binom{5}{4}\binom{54}{1}\binom{1}{1}]/[\binom{59}{5}\binom{35}{1}]$	\$10,000
4. Match four white balls but not Powerball	$[\binom{5}{4}\binom{54}{1}\binom{34}{1}]/[\binom{59}{5}\binom{35}{1}]$	\$100
5. Match three white balls *and* Powerball	$[\binom{5}{3}\binom{54}{2}\binom{1}{1}]/[\binom{59}{5}\binom{35}{1}]$	\$100
6. Match three white balls but not Powerball	$[\binom{5}{3}\binom{54}{2}\binom{34}{1}]/[\binom{59}{5}\binom{35}{1}]$	\$7
7. Match two white balls *and* Powerball	$[\binom{5}{2}\binom{54}{3}\binom{1}{1}]/[\binom{59}{5}\binom{35}{1}]$	\$7
8. Match one white ball *and* Powerball	$[\binom{5}{1}\binom{54}{4}\binom{1}{1}]/[\binom{59}{5}\binom{35}{1}]$	\$4
9. Match zero white balls *and* Powerball	$[\binom{5}{0}\binom{54}{5}\binom{1}{1}]/[\binom{59}{5}\binom{35}{1}]$	\$4

Comment While the six numbers *drawn* in Powerball are random, records show that the six numbers *picked* by ticket-buyers are not random. Why? Because Powerball players have a tendency to bet on their birthdays, which results in their entries tending to have smaller digits and be disproportionately over-represented (and duplicated) in the population of all possible entries. Does that matter? It depends. If one ticket or one hundred thousand tickets qualify for the fourth prize, for example, each of those players will still receive \$100. But if n players qualify for the jackpot, each will receive only l/nth of the jackpot money. So, unless you feel a burning desire to be magnanimous and share your possible (albeit highly improbable) multi-millions of dollars of jackpot winnings with total strangers, betting on your birthday is a bad idea.

Problem-Solving Hints

(Doing combinatorial probability problems)

Listed on p. 70 are several hints that can be helpful in counting the number of ways to do something. Those same hints apply to the solution of combinatorial *probability* problems, but a few others should be kept in mind as well.

(Continued on next page)

(Continued)

1. The solution to a combinatorial probability problem should be set up as a quotient of numerator and denominator *enumerations*. Avoid the temptation to multiply probabilities associated with each position in the sequence. The latter approach will always "sound" reasonable, but it will frequently oversimplify the problem and give the wrong answer.
2. Keep the numerator and denominator consistent with respect to *order*—if permutations are being counted in the numerator, be sure that permutations are being counted in the denominator; likewise, if the outcomes in the numerator are combinations, the outcomes in the denominator must also be combinations.
3. The number of outcomes associated with any problem involving the rolling of n six-sided dice is 6^n; similarly, the number of outcomes associated with tossing a coin n times is 2^n. The number of outcomes associated with dealing a hand of n cards from a standard fifty-two-card poker deck is ${}_{52}C_n$.

Questions

2.7.1. Ten equally qualified marketing assistants are candidates for promotion to associate buyer; seven are men and three are women. If the company intends to promote four of the ten at random, what is the probability that exactly two of the four are women?

2.7.2. An urn contains six chips, numbered 1 through 6. Two are chosen at random and their numbers are added together. What is the probability that the resulting sum is equal to 5?

2.7.3. An urn contains twenty chips, numbered 1 through 20. Two are drawn simultaneously. What is the probability that the numbers on the two chips will differ by more than 2?

2.7.4. A bridge hand (thirteen cards) is dealt from a standard fifty-two-card deck. Let A be the event that the hand contains four aces; let B be the event that the hand contains four kings. Find $P(A \cup B)$.

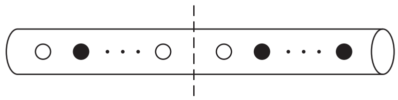
2.7.5. Consider a set of ten urns, nine of which contain three white chips and three red chips each. The tenth contains five white chips and one red chip. An urn is picked at random. Then a sample of size 3 is drawn without replacement from that urn. If all three chips drawn are white, what is the probability that the urn being sampled is the one with five white chips?

2.7.6. A committee of fifty politicians is to be chosen from among our one hundred U.S. senators. If the selection is done at random, what is the probability that each state will be represented?

2.7.7. Suppose that n fair dice are rolled. What are the chances that all n faces will be the same?

2.7.8. Five fair dice are rolled. What is the probability that the faces showing constitute a "full house"—that is, three faces show one number and two faces show a second number?

2.7.9. Imagine that the test tube pictured contains $2n$ grains of sand, n white and n black. Suppose the tube is vigorously shaken. What is the probability that the two colors of sand will completely separate; that is, all of one color fall to the bottom, and all of the other color lie on top? (*Hint:* Consider the $2n$ grains to be aligned in a row. In how many ways can the n white and the n black grains be permuted?)



2.7.10. Does a monkey have a better chance of rearranging

$$ACCLLUUS \quad \text{to spell} \quad CALCULUS$$

or

$$AABEGLR \quad \text{to spell} \quad ALGEBRA?$$

2.7.11. An apartment building has eight floors. If seven people get on the elevator on the first floor, what is the probability they all want to get off on different floors? On the same floor? What assumption are you making? Does it seem reasonable? Explain.

2.7.12. If the letters in the phrase

$$A\ ROLLING\ STONE\ GATHERS\ NO\ MOSS$$

are arranged at random, what are the chances that not all the S's will be adjacent?

2.7.13. Suppose each of ten sticks is broken into a long part and a short part. The twenty parts are arranged into ten pairs and glued back together so that again there are ten sticks. What is the probability that each long part will be paired with a short part? (*Note:* This problem is a model for the effects of radiation on a living cell. Each chromosome, as a result of being struck by ionizing radiation, breaks into two parts, one part containing the centromere. The cell will die unless the fragment containing the centromere recombines with a fragment not containing a centromere.)

2.7.14. Six dice are rolled one time. What is the probability that each of the six faces appears?

2.7.15. Suppose that a randomly selected group of k people are brought together. What is the probability that exactly one pair has the same birthday?

2.7.16. For one-pair poker hands, why is the number of denominations for the three single cards $\binom{12}{3}$ rather than $\binom{12}{1}\binom{11}{1}\binom{10}{1}$?

2.7.17. Dana is not the world's best poker player. Dealt a 2 of diamonds, an 8 of diamonds, an ace of hearts, an ace of clubs, and an ace of spades, she discards the three aces. What are her chances of drawing to a flush?

2.7.18. A poker player is dealt a 7 of diamonds, a queen of diamonds, a queen of hearts, a queen of clubs, and an ace of hearts. He discards the 7. What is his probability of drawing to either a full house or four-of-a-kind?

2.7.19. Tim is dealt a 4 of clubs, a 6 of hearts, an 8 of hearts, a 9 of hearts, and a king of diamonds. He discards the 4 and the king. What are his chances of drawing to a straight flush? To a flush?

2.7.20. Five cards are dealt from a standard 52-card deck. What is the probability that the sum of the faces on the five cards is 48 or more?

2.7.21. Nine cards are dealt from a 52-card deck. Write a formula for the probability that three of the five even numerical denominations are represented twice, one of the three face cards appears twice, and a second face card appears once. (*Note:* Face cards are the jacks, queens, and kings; 2, 4, 6, 8, and 10 are the even numerical denominations.)

2.7.22. A coke hand in bridge is one where none of the thirteen cards is an ace or is higher than a 9. What is the probability of being dealt such a hand?

2.7.23. A pinochle deck has forty-eight cards, two of each of six denominations (9, J, Q, K, 10, A) and the usual four suits. Among the many hands that count for meld is a *roundhouse*, which occurs when a player has a king and queen of each suit. In a hand of twelve cards, what is the probability of getting a "bare" roundhouse (a king and queen of each suit and no other kings or queens)?

2.8 Taking a Second Look at Statistics (Monte Carlo Techniques)

Recall the von Mises definition of probability given on p. 16. If an experiment is repeated n times under identical conditions, and if the event E occurs on m of those repetitions, then

$$P(E) = \lim_{n\to\infty} \frac{m}{n} \tag{2.8.1}$$

To be sure, Equation 2.8.1 is an asymptotic result, but it suggests an obvious (and *very* useful) approximation—if n is finite,

$$P(E) \doteq \frac{m}{n}$$

In general, efforts to estimate probabilities by simulating repetitions of an experiment (usually with a computer) are referred to as *Monte Carlo* studies. Usually the technique is used in situations where an exact probability is difficult to calculate. It can also be used, though, as an empirical justification for choosing one proposed solution over another.

For example, consider the game described in Example 2.4.10. An urn contains a red chip, a blue chip, and a two-color chip (red on one side, blue on the other). One chip is drawn at random and placed on a table. The question is, if *blue* is showing, what is the probability that the color underneath is also *blue*?

Pictured in Figure 2.8.1 are two ways of conceptualizing the question just posed. The outcomes in (a) are assuming that a *chip* was drawn. Starting with that premise, the answer to the question is $\frac{1}{2}$—the red chip is obviously eliminated and only one of the two remaining chips is blue on both sides.

$$\begin{array}{c}\underline{\text{Chip drawn}}\\ \left.\begin{array}{c}\text{red}\\ \text{blue}\\ \text{two-color}\end{array}\right\}\longrightarrow P(B|B)=1/2\\ \textbf{(a)}\end{array}\qquad\begin{array}{c}\underline{\text{Side drawn}}\\ \left.\begin{array}{c}\text{red/red}\\ \text{blue/blue}\\ \text{red/blue}\end{array}\right\}\longrightarrow P(B|B)=2/3\\ \textbf{(b)}\end{array}$$

Figure 2.8.1

By way of contrast, the outcomes in (b) are assuming that the *side* of a chip was drawn. If so, the blue color showing could be any of three blue sides, two of which are blue underneath. According to model (b), then, the probability of both sides being blue is $\frac{2}{3}$.

The formal analysis on p. 43, of course, resolves the debate—the correct answer is $\frac{2}{3}$. But suppose that such a derivation was unavailable. How might we assess the relative plausibilities of $\frac{1}{2}$ and $\frac{2}{3}$? The answer is simple—just play the game a number of times and see what proportion of outcomes that show blue on top have blue underneath.

To that end, Table 2.8.1 summarizes the results of one hundred random drawings. For a total of fifty-two trials, blue was showing (S) when the chip was placed on a

Table 2.8.1

Trial #	S	U	Trial #	S	U	Trial #	S	U	Trial #	S	U
1	R	B	26	B	R	51	B	R	76	B	B*
2	B	B*	27	R	R	52	R	B	77	B	B*
3	B	R	28	R	B	53	B	B*	78	R	R
4	R	R	29	R	B	54	R	B	79	B	B*
5	R	B	30	R	R	55	R	R	80	R	R
6	R	B	31	R	B	56	R	B	81	R	B
7	R	R	32	B	B*	57	R	R	82	R	B
8	R	R	33	R	B	58	B	B*	83	R	R
9	B	B*	34	B	B*	59	B	R	84	B	R
10	B	R	35	B	B*	60	B	B*	85	B	R
11	R	R	36	R	R	61	B	R	86	R	R
12	B	B*	37	B	R	62	R	B	87	B	B*
13	R	R	38	B	B*	63	R	R	88	R	B
14	B	R	39	R	R	64	R	R	89	B	R
15	B	B*	40	B	B*	65	B	B*	90	R	R
16	B	B*	41	B	B*	66	B	R	91	R	B
17	R	B	42	B	R	67	R	R	92	R	R
18	B	R	43	B	B*	68	B	B*	93	R	R
19	B	B*	44	B	B*	69	B	B*	94	R	B
20	B	B*	45	B	B*	70	R	R	95	B	B*
21	R	R	46	R	R	71	R	R	96	B	B*
22	R	R	47	B	B*	72	B	B*	97	B	R
23	B	B*	48	B	B*	73	R	B	98	R	R
24	B	R	49	R	R	74	R	R	99	B	B*
25	B	B*	50	R	R	75	B	B*	100	B	B*

table; for thirty-six of the trials (those marked with an asterisk), the color underneath (U) was also blue. Using the approximation suggested by Equation 2.8.1,

$$P(\text{Blue is underneath} \mid \text{Blue is on top}) = P(B \mid B) \doteq \frac{36}{52} = 0.69$$

a figure much more consistent with $\frac{2}{3}$ than with $\frac{1}{2}$.

The point of this example is not to downgrade the importance of rigorous derivations and exact answers. Far from it. The application of Theorem 2.4.1 to solve the problem posed in Example 2.4.10 is obviously superior to the Monte Carlo approximation illustrated in Table 2.8.1. Still, replications of experiments can often provide valuable insights and call attention to nuances that might otherwise go unnoticed.