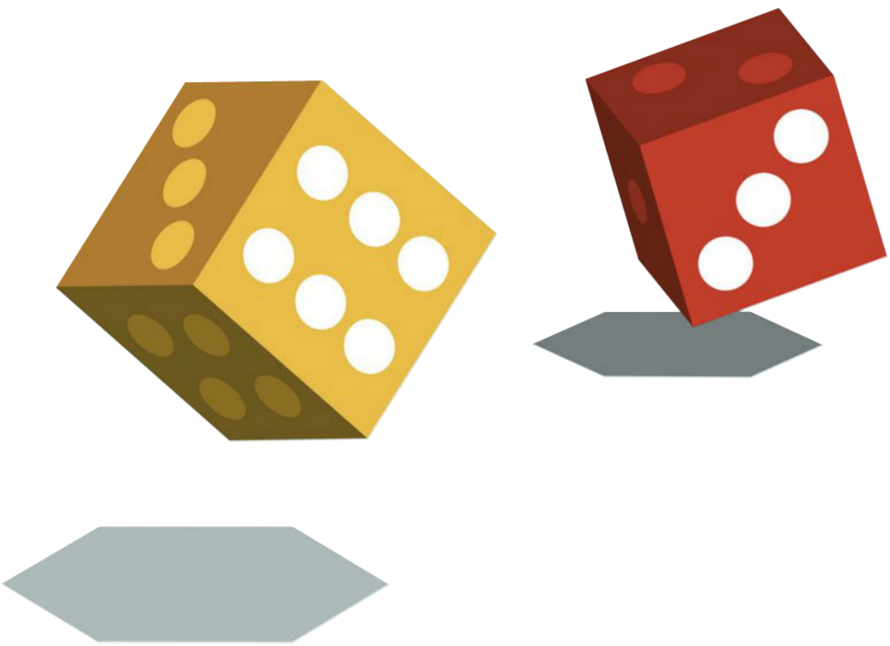
15 Chances, Probabilities, and Odds

Measuring Uncertainty

With the possible exception of death and taxes, pretty much everything else that happens to us in our lives is layered with some degree of *uncertainty*. That's why we pay attention to the weather report, buy insurance, and constantly wish our friends "good luck."

Although we are all familiar with uncertainty, we don't always have a good grasp on how to measure it. Some situations involve only a small degree of uncertainty ("I'm pretty sure I aced the midterm"), some situations involve a large degree of uncertainty ("I have no clue how I did in that midterm"), and some situations fall in between ("I think I did well in the midterm, but . . ."). To quantify and measure more precisely the amount of uncertainty in the many uncertain events that affect our lives we use the concept of *probability*.



“Nothing in life is certain except death and taxes.”
—Ben Franklin

Chance, probability, odds—we use these words carelessly in casual conversation, and most of the time we can get away with it. Technically speaking, however, each of these words represents a slightly different way to measure the *likelihood* of an event. When we speak of the *chance* of some event happening we express it in terms of percentages, such as when the weatherperson reports that "the chance of rain tomorrow is 40%." When we speak of the *probability* of some event happening we express it in terms of a ratio ("the probability is 2 out of 5") or, equivalently as a fraction $\left(\frac{2}{5}\right)$ or a decimal (0.4). Finally, when we speak of the *odds in favor* of an event, we use a pair of numbers, as in "the odds of the Red Sox winning the World Series are 2 to 3."

In this chapter we will learn how to interpret and work with probabilities, chances, and odds in a formal mathematical context. This will be our very brief introduction to the mathematical theory of probability, a relatively young branch of mathematics that has become critically important to many aspects of modern life. Insurance, public health, science, sports, gambling, the stock market—wherever there is uncertainty to be tamed—the mathematical theory of probability plays a significant role.

Our discussion in this chapter is divided into two parts. In the first part we introduce the basic theoretical framework needed for a meaningful discussion of probabilities: the concepts of *random experiment* and *sample space* (Section 15.1), the basic rules of *counting* (Section 15.2), and the dual concepts of *permutation* and *combination* (Section 15.3). In the second part of the chapter we discuss general *probability spaces* (Section 15.4) and probabilities in spaces in which all outcomes are equally likely (Section 15.5). The chapter concludes with a brief discussion of *odds* and their relationship to probabilities (Section 15.6).

15.1 Random Experiments and Sample Spaces

In broad terms, probability is the *quantification of uncertainty*. To understand what that means, we may start by formalizing the notion of uncertainty.

We will use the term **random experiment** to describe an activity or a process *whose outcome cannot be predicted ahead of time*. Typical examples of random experiments are tossing a coin, rolling a pair of dice, drawing cards out of a deck of cards, predicting the result of a football game, and forecasting the path of a hurricane.

Associated with every random experiment is the *set* of all of its possible outcomes, called the **sample space** of the experiment. For the sake of simplicity, we will concentrate on experiments for which there is only a finite set of outcomes, although experiments with infinitely many outcomes are both possible and important.

We illustrate the importance of the sample space by means of several examples. Since the sample space of any experiment is a set of outcomes, we will use set notation to describe it. We will consistently use the letter S to denote a sample space and the letter N to denote the *size* of the sample space S (i.e., the number of outcomes in S).

EXAMPLE 15.1 Tossing a Coin

One simple random experiment is to *toss a quarter* and *observe whether it lands heads or tails*. The sample space can be described by $S = \{H, T\}$, where H stands for *Heads* and T for *Tails*. Here $N = 2$.

A couple of comments about coins are in order here. First, the fact that the coin in Example 15.1 was a quarter is essentially irrelevant. Practically all coins have an obvious "heads" side (and thus a "tails" side), and even when they don't—as in a "buffalo nickel"—we can agree ahead of time which side is which. Second, there are fake coins out there on which both sides are "heads." Tossing such a coin does not fit our definition of a random experiment, so from now on, we will assume that all coins used in our experiments have two different sides, which we will call H and T.

EXAMPLE 15.2 More Coin Tossing

Suppose we toss a coin *twice* and *record* the outcome of each toss (H or T) in the order it happens. The sample space now is $S = \{HH, HT, TH, TT\}$, where HT means that the first toss came up H and the second toss came up T, which is a different outcome from TH (first toss T and second toss H). In this sample space $N = 4$.

Suppose now we *toss two distinguishable coins* (say, a nickel and a quarter) *at the same time* (tricky but definitely possible). This random experiment appears different from the one where we toss one coin twice, but the sample space is still $S = \{HH, HT, TH, TT\}$. (Here we must agree what the order of the symbols is—for example, the first symbol describes the quarter and the second the nickel.)

Since they have the same sample space, we will consider the two random experiments just described as the same random experiment.

Now let's consider a different random experiment. We are still tossing a coin twice, but we only care now about the *number of heads* that come up. Here there are only three possible outcomes (no heads, one head, or both heads), and symbolically we might describe this sample space as $S = \{0, 1, 2\}$.

The important point made in Example 15.2 is that a random experiment is defined by two things: the action (such as tossing coins) and what it is that we are interested in observing from the action.

EXAMPLE 15.3 Shooting Free Throws

Here is a familiar scenario: Your favorite basketball team is down by 1, clock running out, and one of your players is fouled and goes to the line to shoot free throws, with the game riding on the outcome. It's not a good time to think of sample spaces, but let's do it anyway.

Clearly the shooting of free throws is a random experiment, but what is the sample space? As in Example 15.2, the answer depends on a few subtleties.

In one scenario (the *penalty situation*) your player is going to shoot two free throws no matter what. In this case one could argue that what really matters is how many free throws he or she makes (make both and win the game, miss one and tie and go to overtime, miss both and lose the game). When we look at it this way the sample space is $S = \{0, 1, 2\}$.

A somewhat more stressful scenario is when your player is shooting a *one-and-one*. This means that the player gets to shoot the second free throw only if he or she makes the first one. In this case there are also three possible outcomes, but the circumstances are different because the order of events is relevant (miss the first free throw and lose the game, make the first free throw but miss the second one and tie the game, make both and win the game). We can describe this sample space as $S = \{f, sf, ss\}$, where we use f to indicate *failure* (missed the free throw) and s to indicate *success*.

We will now discuss a couple of examples of random experiments involving dice. A die is a cube, usually made of plastic, whose six faces are marked with dots (from 1 to 6) called "pips." Random experiments using dice have a long-standing tradition in our culture and are a part of both gambling and recreational games such as Monopoly or Yahtzee.

EXAMPLE 15.4 Rolling a Pair of Dice

The most common scenario when rolling a pair of dice is to only consider the *total* of the two numbers rolled. In this situation we don't really care how a particular total comes about. We can "roll a seven" in various paired combinations—a 3 and a 4, a 2 and a 5, a 1 and a 6. No matter how the individual dice come up, the only thing that matters is the total rolled.

The possible outcomes in this scenario range from "rolling a two" to "rolling a twelve," and the sample space can be described by $S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.

EXAMPLE 15.5 More Dice Rolling

A more general scenario when rolling a pair of dice is when we do care what number each individual die turns up (in certain bets in craps, for example, two

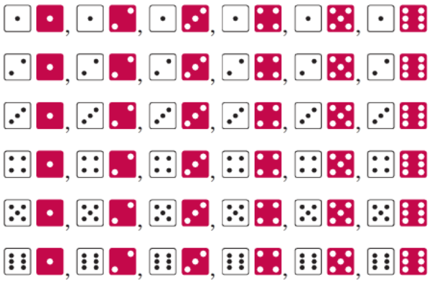


FIGURE 15-1

"fours" is a winner but a "five" and a "three" is not). Here we have a sample space with 36 different outcomes, shown in Fig. 15-1. Notice that in Fig. 15-1 the dice are colored white and red, a symbolic way to emphasize the fact that we are treating the dice as *distinguishable* objects. This is why the rolls ⚁⚄ and ⚄⚁ are considered different outcomes.

EXAMPLE 15.6 **Ranking the Candidates in an Election**

Five candidates (A, B, C, D, and E) are running in an election. The top three vote getters are chosen President, Vice President, and Secretary, in that order. Such an election can be considered a random experiment with sample space $S = \{ABC, ACB, BAC, BCA, CAB, CBA, ABD, ADB, \ldots, CDE\}$. We will assume that outcome ABC denotes that candidate A is elected President, B is elected Vice President, and C is elected Secretary. This is important because the outcomes ABC and BAC are different outcomes and must appear separately in the sample space.

Because this sample space is a bit too large to write out in full (we will soon learn that its size is $N = 60$), we use the ". . ." notation. It is a way of saying "and so on—you get the point."

Example 15.6 illustrates the point that sample spaces can have a lot of outcomes and that we are often justified in not writing each and every one of them down. This is where the ". . ." comes in handy. The key thing is to understand what the sample space looks like without necessarily writing all the outcomes down. Our real goal is to find N, the size of the sample space. If we can do it without having to list all the outcomes, then so much the better. We will discuss how this is done in the next two sections.

15.2 Counting Outcomes in Sample Spaces

EXAMPLE 15.7 **Tossing More Coins**

If we *toss a coin three times* and separately record the outcome of each toss, the sample space is given by $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$. Here we can just count the outcomes and get $N = 8$.

Now let's look at an expanded version of the same idea and *toss a coin 10 times* (and record the outcome of each toss). In this case the sample space S is too big to write down, but we can "count" the number of outcomes in S without having to tally them one by one. Here is how we do it.

Each outcome in S can be described by a string of 10 consecutive letters, where each letter is either H or T. (This is a natural extension of the ideas introduced in Example 15.2.) For example, $THHTHTHHTT$ represents a *single* outcome in our sample space—the one in which the first toss came up T, the second toss came up H, the third toss came up H, and so on. To count all the outcomes, we will argue as follows: There are two choices for the first letter (H or T), two choices for the second letter, two choices for the third letter, and so on down to the tenth letter. The total number of possible strings is found by *multiplying* the number of choices for each letter. Thus $N = 2^{10} = 1024$.

The rule we used in Example 15.6 is an old friend from Chapter 2—the **multiplication rule**, and for a formal definition of the multiplication rule the reader is referred to Section 2.4. Informally, the multiplication rule simply says that *when something is done in stages, the number of ways it can be done is found by multiplying the number of ways each of the stages can be done*. The easiest way to understand this simple but powerful idea is by looking at some examples.

EXAMPLE 15.8 The Making of a Wardrobe

Dolores is a young saleswoman planning her next business trip. She is thinking about packing three different pairs of shoes, four skirts, six blouses, and two jackets. If all the items are color coordinated, how many different *outfits* will she be able to create by combining these items?

To answer this question, we must first define what we mean by an "outfit." Let's assume that an outfit consists of one pair of shoes, one skirt, one blouse, and one jacket. Then to make an outfit Dolores must choose a pair of shoes (three choices), a skirt (four choices), a blouse (six choices), and a jacket (two choices). By the multiplication rule the total number of possible outfits is $3 \times 4 \times 6 \times 2 = 144$. (Think about it—Dolores can be on the road for over four months and never have to wear the same outfit twice! And it all fits in a small suitcase.)

EXAMPLE 15.9 The Making of a Wardrobe: Part 2

Once again, Dolores is packing for a business trip. This time, she packs three pairs of shoes, four skirts, three pairs of slacks, six blouses, three turtlenecks, and two jackets. As before, we can assume that she coordinates the colors so that everything goes with everything else. This time, we will define an outfit as consisting of a pair of shoes, a choice of "lower wear" (either a skirt *or* a pair of slacks), and a choice of "upper wear" (it could be a blouse *or* a turtleneck *or both*), and, finally, she may or may not choose to wear a jacket. How many different such outfits are possible?

This is a more sophisticated variation of Example 15.8. Our strategy will be to think of an outfit as being put together in stages and to draw a box for each of the stages. We then separately count the number of choices at each stage and enter that number in the corresponding box. (Some of these calculations can themselves be mini-counting problems.) The last step is to multiply the numbers in each box. The details are illustrated in Fig. 15-2. The final count for the number of different outfits is $N = 3 \times 7 \times 27 \times 3 = 1701$.

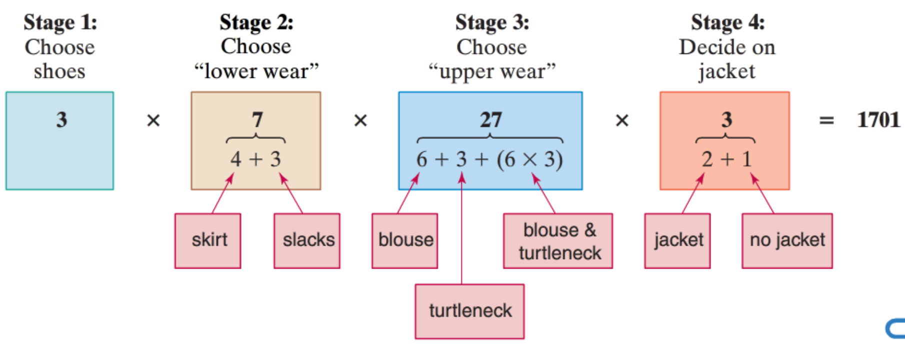


FIGURE 15-2

The method of drawing boxes representing the successive stages in a process, and putting the number of choices for each stage inside the box is a convenient strategy that often helps clarify one's thinking. Silly as it may seem, we strongly recommend it. For ease of reference, we will call it the *box model* for counting.

EXAMPLE 15.10 **Ranking the Candidates in an Election: Part 2**

This example is a follow-up to Example 15.6. Five candidates are running in an election, with the top three vote getters elected (in order) as President, Vice President, and Secretary. We want to know how big the sample space is. Using a box model, we see that this becomes a reasonably easy counting problem, as illustrated in Fig. 15-3.

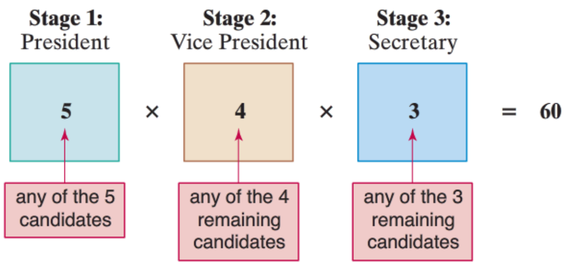


FIGURE 15-3

15.3 Permutations and Combinations

Many counting problems can be reduced to a question of counting the number of ways in which we can choose groups of objects from a larger group of objects. Often these problems require somewhat more sophisticated counting methods than the plain vanilla multiplication rule. In this section we will discuss the dual concepts of *permutation* (a group of objects in which the ordering of the objects within the group *makes a difference*) and *combination* (a group of objects in which the ordering of the objects is *irrelevant*).

EXAMPLE 15.11 **The Pleasures of Ice Cream**

Baskin-Robbins offers 31 different flavors of ice cream. A "true double" is the name we will use for two scoops of ice cream of two *different* flavors. Say you want a true double in a bowl—how many different choices do you have?

The natural impulse is to count the number of choices using the multiplication rule (and a box model) as shown in Fig. 15-4. This would give an answer of $31 \times 30 = 930$ true doubles. Unfortunately, this answer is *double counting* each of the true doubles. Why?

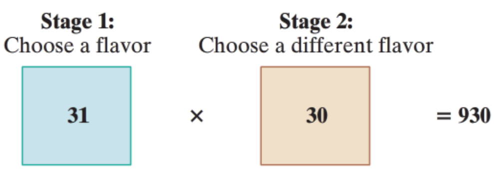


FIGURE 15-4

When we use the multiplication rule, *there is a well-defined order to things*, and a scoop of strawberry followed by a scoop of chocolate is counted separately from a scoop of chocolate followed by a scoop of strawberry. But by all reasonable standards, a bowl of chocolate-strawberry is the same as a bowl of strawberry-chocolate. (This is not necessarily true when the ice cream is served in a cone. Fussy people can be very picky about the order of the flavors when the scoops are stacked up.)

The good news is that now that we understand why the count of 930 is wrong we can fix it. All we have to do to get the correct answer is to divide the original count by 2. It follows that the number of true double choices at Baskin-Robbins is $(31 \times 30)/2 = 465$.

Example 15.11 is an important one. It warns us that we have to be careful about how we use the multiplication rule and box models in counting problems where the order in which we choose the objects (ice cream flavors) does not affect the answer. Let's take this idea to the next level.

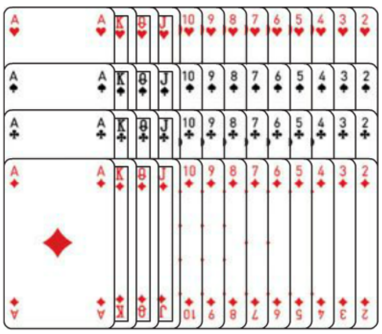
EXAMPLE 15.12 The Pleasures of Ice Cream: Part 2

Some days you have a real craving for ice cream, and on such days you like to go to Baskin-Robbins and order a *true triple* (in a bowl). How many different choices do you have? (As you might have guessed, a "true triple" consists of three scoops of ice cream each of a different flavor.)

Starting with the multiplication rule, we have 31 choices for the "first" flavor, 30 choices for the "second" flavor, and 29 choices for the "third" flavor, for an apparent grand total of $31 \times 30 \times 29 = 26{,}970$. But once again this answer counts each true triple more than once; in fact, it does so *six times*! (More on that shortly.) If we accept this, the correct answer must be $26{,}970/6 = 4495$.

Why is it that the count of 26,970 counts each true triple *six* times? The answer comes down to this: Any three flavors (call them X, Y, and Z) can be listed in $3 \times 2 \times 1 = 6$ different ways (XYZ, XZY, YXZ, YZX, ZXY, and ZYX). The multiplication rule counts each of these separately, but when you think in terms of ice cream scoops in a bowl, they are the same true triple regardless of the order.

The bottom line is that there are 4495 different possibilities for a true triple at Baskin-Robbins. We can better understand where this number comes from by looking at it in its raw, uncalculated form $[(31 \times 30 \times 29)/(3 \times 2 \times 1) = 4495]$. The numerator $(31 \times 30 \times 29)$ comes from counting *ordered* triples using the multiplication rule; the denominator $(3 \times 2 \times 1)$ comes from counting the number of ways in which three things (in this case, the three flavors in a triple) can be rearranged. The denominator $3 \times 2 \times 1$ is already familiar to us—it is the *factorial* of 3. (We discussed the factorial in Chapters 2 and 6, so we won't dwell on it here.)



Our next example will deal with the game of poker. Despite the great deal of exposure poker gets on television these days, there are a lot of misconceptions about the mathematics behind poker hands. For readers not familiar with the game, poker is a betting game played with a standard deck of cards [a standard deck of cards has 52 cards divided into 4 *suits* (clubs, diamonds, hearts, and spades) and with 13 *values* in each suit (2, 3, . . ., 10, J, Q, K, A)]. There are many variations of poker depending on how many cards are dealt and whether all the cards are *down* cards (only the player receiving the card can see it) or some of the cards are *up* cards (dealt face up so that they can be seen by all the players).

EXAMPLE 15.13 Five-Card Poker Hands

In this example we will compare two types of games: five-card *stud poker* and five-card *draw poker*. In both of these games a player ends up with five cards, but there is an important difference when analyzing the mathematics behind the games: In five-card *draw* the order in which the cards come up is irrelevant; in five-card *stud* the order in which the cards come up is extremely relevant. The reason for this is that in five-card draw all cards are dealt down, but in five-card stud only the first card is dealt down—the remaining four cards are dealt up, one at a time. This means that players can assess the relative strengths of the other players' hands as the game progresses and play their hands accordingly.

Left: Five-card *draw* poker hand: The order in which the cards are dealt is irrelevant. *Center and right*: Five-card *stud* poker hands. The order of the up cards is important. The hand on the right is better than the hand in the center.

Counting the number of five-card *stud* poker hands is a direct application of the multiplication rule: 52 possibilities for the first card, 51 for the second card, 50 for the third card, 49 for the fourth card, and 48 for the fifth card, for an awesome total of $52 \times 51 \times 50 \times 49 \times 48 = 311{,}875{,}200$ possible hands.

Counting the number of five-card *draw* poker hands requires a little more finesse. Here a player gets five down cards and the hand is the same regardless of the order in which the cards are dealt. There are $5! = 120$ different ways in which the same set of five cards can be ordered, so that one draw hand corresponds to 120 different stud hands. Thus, the stud hands count is exactly 120 times bigger than the draw hands count. Great! All we have to do then is divide the 311,875,200 (number of stud hands) by 120 and get our answer: There are 2,598,960 possible five-card draw hands. [As before, it's more telling to look at this answer in the uncalculated form $(52 \times 51 \times 50 \times 49 \times 48)/5! = 2{,}598{,}960$.]

We are now ready to generalize the ideas developed in Examples 15.12 and 15.13. Suppose that we have a set of n distinct objects and we want to select r different objects from this set. The number of ways that this can be done depends on whether the selections are *ordered* or *unordered*. Ordered selections are the generalization of stud poker hands—selecting the same objects in different order gets you something different. Unordered selections are the generalization of draw poker hands—selecting the same objects in different order gets you nothing new. To distinguish between these two scenarios, we use the terms **permutation** to describe an ordered selection and **combination** to describe an unordered selection. (One way to remember which is which is to remember that there are many more permutations than there are combinations.)

■ You should take a look at your calculator and figure out the correct sequence of keystrokes to compute these numbers, and then try Exercises 31 and 32.

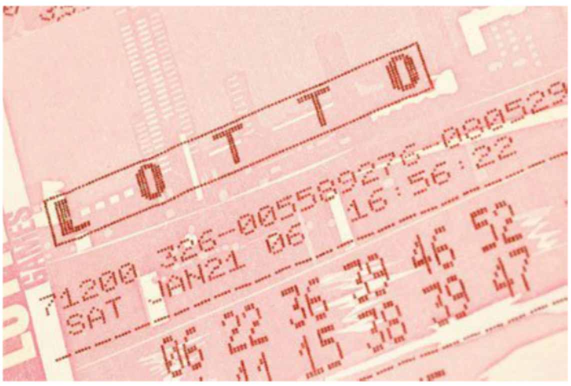
For a given number of objects n and a given selection size r (where $0 \le r \le n$), we can talk about the "number of permutations of n objects taken r at a time" and the "number of combinations of n objects taken r at a time," and these two extremely important families of numbers are denoted ${}_nP_r$ and ${}_nC_r$, respectively. (Some calculators use variations of this notation, such as $P_{n,r}$ and $C_{n,r}$, respectively.)

A summary of the essential facts about the numbers ${}_nP_r$ and ${}_nC_r$ is given in Table 15-1.

TABLE 15-1 Permutations and Combinations

Notation	${}_nP_r$	${}_nC_r$
Formula 1	${}_nP_r = n(n-1)(n-2)\cdots(n-r+1)$	${}_nC_r = \dfrac{n(n-1)(n-2)\cdots(n-r+1)}{r!}$
Formula 2	${}_nP_r = \dfrac{n!}{(n-r)!}$	${}_nC_r = \dfrac{n!}{(n-r)!r!}$
Applications	Stud poker hands, rankings, committees with assignments	Draw poker hands, lottery tickets, coalitions, subsets

EXAMPLE 15.14 **The Florida Lotto**



Like many other state lotteries, the Florida Lotto is a game in which for a small investment of just one dollar a player has a chance of winning tens of millions of dollars. Enormous upside, hardly any downside—that's why people love playing the lottery and, like they say, "everybody has to have a dream." But, in general, lotteries are a very bad investment, even if it's only a dollar, and the dreams can turn to nightmares. Why so?

In a Florida Lotto ticket, one gets to select *six* numbers from 1 through 53. To win the jackpot (there are other lesser prizes we won't discuss here), those six numbers have to match the winning numbers drawn by the lottery in any order. Since a lottery draw is just an unordered selection of six objects (the winning numbers) out of 53 objects (the numbers 1 through 53), the number of possible draws is ${}_{53}C_6$ (see Table 15-1). Doesn't sound too bad until we do the computation (or use a calculator) and realize that

$${}_{53}C_6 = \frac{53 \times 52 \times 51 \times 50 \times 49 \times 48}{6!} = 22{,}957{,}480.$$

15.4 Probability Spaces

What Is a Probability?

If we toss a coin in the air, *what is the probability that it will land heads up*? This one is not a very profound question, and almost everybody agrees on the answer, although not necessarily for the same reason. The standard answer given is 1 out of 2, or 1/2. But why is the answer 1/2, and what does such an answer mean?

One common explanation given for the answer of 1/2 is that when we toss a coin, there are two possible outcomes (H and T), and since H represents one of the two possibilities, the probability of the outcome H must be 1 out of 2 or 1/2. This logic, while correct in the case of an honest coin, has a lot of holes in it.

Consider how the same argument would sound in a different scenario.

EXAMPLE 15.15 **Who Is Shooting Those Free Throws?**

Imagine an NBA player in the act of shooting a free throw. Just like with a coin toss, there are two possible outcomes to the free-throw shot (success or failure), but it would be absurd to conclude that the probability of making the free throw is therefore 1 out of 2, or 1/2. Here the two outcomes are not both equally likely, and their probabilities should reflect that.

The probability of a basketball player making a free throw very much depends on the abilities of the player doing the shooting—it makes a difference if it's Steve Nash or Shaquille O'Neal. Nash is one of the best free-throw shooters in the history of the NBA, with a career average of 90%, while Shaq is a notoriously poor free-throw shooter (52% career average). These percentages, over the long term, represent an approximation of the true probability each of them has of making a free throw—about 0.90 and 0.52, respectively. In either case, the probability is not 0.5.

Example 15.15 leads us to what is known as the *empirical* interpretation of the concept of probability. Under this interpretation when tossing an honest coin the probability of *Heads* is 1/2 not because *Heads* is one out of two possible outcomes but because, if we were to toss the coin over and over—hundreds, possibly thousands, of times—in the long run about half of the tosses will turn out to be heads, a fact that has been confirmed by experiment many times.

The argument as to exactly how to interpret the statement "the probability of X is such and such" goes back to the late 1600s, and it wasn't until the 1930s that a formal theory for dealing with probabilities was developed by the Russian mathematician A. N. Kolmogorov (1903–1987). This theory has made probability one of the most useful and important concepts in modern mathematics. In the remainder of this chapter we will discuss some of the basic concepts of *probability theory*.

Events

An **event** is any subset of the sample space. That is, an event is any set of individual outcomes. (This definition includes the possibility of an "event" that has no outcomes as well as events consisting of a single outcome.) By definition, events are sets (subsets of the sample space), and we will deal with events using set notation as well as the basic rules of set theory.

A convenient way to think of an event is as a package in which outcomes with some common characteristic are bundled together. Say you are rolling a pair of dice and hoping that on the next roll "you roll a 7"—that would make you temporarily rich. The best way to describe what really matters (to you) is by packaging together all the different ways to "roll a 7" as a single event E:

E = {⚀⚅, ⚁⚄, ⚂⚃, ⚃⚂, ⚄⚁, ⚅⚀}

Sometimes an event consists of just one outcome. We will call such an event a **simple event**. In some sense simple events are the building blocks for all other events (more on that later). There is also the special case of the empty set { }, corresponding to an event with no outcomes. Such an event can never happen, and thus we call it the **impossible event**.

EXAMPLE 15.16 **Coin-Tossing Events**

Let's revisit the experiment of tossing a coin three times and recording the result of each toss (Example 15.7). The sample space for this experiment is $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$. The set S has hundreds of subsets (256 to be exact), each of which represents a different event. Table 15-2 shows just a few of these events.

TABLE 15-2

Event (in words)	Event (as a set)
Toss 2 or more heads.	$\{HHT, HTH, THH, HHH\}$
Toss more than 2 heads.	$\{HHH\}$
Toss 2 heads or fewer.	$\{TTT, TTH, THT, HTT, THH, HTH, HHT\}$
Toss no tails.	$\{HHH\}$
Toss exactly 1 tail.	$\{HHT, HTH, THH\}$
Toss exactly 1 head.	$\{HTT, THT, TTH\}$
First toss is heads.	$\{HHH, HHT, HTH, HTT\}$
Toss same number of heads as tails.	$\{\ \}$ (*Note*: This is the *impossible event*.)
Toss 3 heads or fewer.	S (*Note*: This event is called the *certain event*.)

Probability Assignments

Let's return to free-throw shooting, as it is a useful metaphor for many probability questions.

EXAMPLE 15.17 **The Unknown Free-Throw Shooter**

A player is going to shoot a free throw. We know nothing about his or her abilities—for all we know, the player could be Steve Nash, or Shaquille O'Neal, or Joe Schmoe, or you, or me.

How can we describe the probability that he or she will make that free throw? It seems that there is no way to answer this question, since we know nothing about the ability of the shooter. We could argue that the probability could be just about any number between 0 and 1. No problem—we make our unknown probability a variable, say p.

What can we say about the probability that our shooter misses the free throw? A lot. Since there are only two possible outcomes in the sample space $S = \{s, f\}$, the probability of success (s) and the probability of failure (f) must complement each other—in other words, must add up to 1. This means that the probability of missing the free throw must be $1 - p$.

Table 15-3 is a summary of the line on a generic free-throw shooter. Humble as it may seem, Table 15-3 gives a complete model of free-throw shooting. It works when the free-throw shooter is Steve Nash (make it $p = 0.90$), Shaquille O'Neal (make it $p = 0.52$), or the author of this book (make it $p = 0.30$). Each one of the choices results in a different assignment of numbers to the outcomes in the sample space.

TABLE 15-3

Event	Probability
$\{\ \}$	0
$\{s\}$	p
$\{f\}$	$1 - p$
$\{s, f\}$	1

Example 15.17 illustrates the concept of a *probability assignment*. A **probability assignment** is a function that assigns to each event E a number between

0 and 1, which represents the probability of the event E and which we denote by $\Pr(E)$. A probability assignment always assigns probability 0 to the *impossible event* $[\Pr(\{\ \}) = 0]$ and probability 1 to the whole sample space $[\Pr(S) = 1]$.

With finite sample spaces a probability assignment is defined by assigning probabilities to just the simple events in the sample space. Once we do this, we can find the probability of any event by simply *adding the probabilities of the individual outcomes that make up that event*. There are only two requirements for a valid probability assignment: (1) *All probabilities are numbers between 0 and 1*, and (2) *the sum of the probabilities of the simple events equals 1*.

■ In the case of a simple event $\{a\}$ we cheat a little and for the sake of simplicity we speak of "the probability of the outcome a" when we really should say "the probability of the event $\{a\}$." Accordingly, we will write $\Pr(a)$ in lieu of the technically correct but awkward $\Pr(\{a\})$.

When bookmakers or professional odds-makers handicap a sporting event, they do so by essentially defining a probability assignment for the sample space of all possible outcomes of that event. The next example is a simple illustration of how this might be done.

EXAMPLE 15.18 **Handicapping a Tennis Tournament**

There are six players playing in a tennis tournament: A (Russian, female), B (Croatian, male), C (Australian, male), D (Swiss, male), E (American, female), and F (American, female).

To handicap the winner of the tournament we need a probability assignment on the sample space $S = \{A, B, C, D, E, F\}$. With sporting events the probability assignment is subjective (it reflects an opinion), but professional odds-makers are usually very good at getting close to the right probabilities. For example, imagine that a professional odds-maker comes up with the following probability assignment: $\Pr(A) = 0.08$, $\Pr(B) = 0.16$, $\Pr(C) = 0.20$, $\Pr(D) = 0.25$, $\Pr(E) = 0.16$. [We are missing $\Pr(F)$ from the list, but since the probabilities of the simple events must add up to 1, we can do the arithmetic: $\Pr(F) = 0.15$.]

Once we have the probabilities of the simple events, the probabilities of all other events follow by addition. For example, the probability that an American will win the tournament is given by $\Pr(E) + \Pr(F) = 0.16 + 0.15 = 0.31$. Likewise, the probability that a male will win the tournament is given by $\Pr(B) + \Pr(C) + \Pr(D) = 0.16 + 0.20 + 0.25 = 0.61$. The probability that an American male will win the tournament is $\Pr(\{\ \}) = 0$, since this one is an impossible event—there are no American males in the tournament!

The probability assignment discussed in Example 15.18 reflects the opinion of one specific observer. A different odds-maker might have a slightly different perspective and come up with a different probability assignment. This underscores the fact that sometimes there is no one single "correct" probability assignment on a sample space.

Once a specific probability assignment is made on a sample space, the combination of the sample space and the probability assignment is called a **probability space**. The following is a summary of the key facts related to probability spaces.

■ In this chapter the term *probability space* will always refer to a *finite* probability space. Infinite probability spaces are important and interesting but require a much higher level of mathematics and we will not discuss them here.

ELEMENTS OF A PROBABILITY SPACE

- **Sample space:** $S = \{o_1, o_2, \ldots, o_N\}$.
- **Probability assignment:** $\Pr(o_1), \Pr(o_2), \ldots, \Pr(o_N)$.
  [Each of these is a number between 0 and 1 satisfying
  $\Pr(o_1) + \Pr(o_2) + \cdots + \Pr(o_N) = 1$.]
- **Events:** These are all the subsets of S, including $\{\ \}$ and S itself. The probability of an event is given by the sum of the probabilities of the individual outcomes that make up the event. [In particular, $\Pr(\{\ \}) = 0$ and $\Pr(S) = 1$.]

15.5 Equiprobable Spaces

One of the most common uses of randomness in the real world is as a mechanism to guarantee fairness. That's why it is universally accepted that tossing a coin, rolling a die, or drawing cards is a fair way of choosing among equally deserving choices. This is true as long as the coin, die, or deck of cards is "honest."

What does *honesty* mean when applied to coins, dice, or decks of cards? It essentially means that all individual outcomes in the sample space are equally probable. Thus, an *honest* coin is one in which H and T have the same probability of coming up, and an *honest* die is one in which each of the numbers 1 through 6 is equally likely to be rolled. A probability space in which each simple event has an equal probability is called an **equiprobable space**. (Informally, you can think of an equiprobable space as an "equal opportunity" probability space.)

In equiprobable spaces, calculating probabilities of events becomes simply a matter of counting. First, we need to find N, the size of the sample space. Each individual outcome in the sample space will have probability equal to $1/N$ (they all have the same probability, and the sum of the probabilities must equal 1). To find the probability of the event E we then find k, the number of outcomes in E. Since each of these outcomes has probability $1/N$, the probability of E is then k/N.

PROBABILITIES IN EQUIPROBABLE SPACES

If k denotes the size of an event E and N denotes the size of the sample space S, then in an equiprobable space

$$\Pr(E) = \frac{k}{N}$$

EXAMPLE 15.19 **Honest Coin Tossing**

This example is a follow-up of Example 15.16. Suppose that a coin is tossed three times, and we have been assured that the coin is an honest coin. If this is true, then each of the eight possible outcomes in the sample space has probability $1/8$. (Recall that there are $N = 8$ outcomes in the sample space.)

The probability of any event E is given by the number of outcomes in E divided by 8. Table 15-4 shows each of the events in Table 15-2 with their respective probabilities.

TABLE 15-4

Event (in words)	Event (as a set)	Probability
Toss 2 or more heads.	$\{HHT, HTH, THH, HHH\}$	$4/8 = 1/2$
Toss more than 2 heads.	$\{HHH\}$	$1/8$
Toss 2 heads or fewer.	$\{TTT, TTH, THT, HTT, THH, HTH, HHT\}$	$7/8$
Toss no tails.	$\{HHH\}$	$1/8$
Toss exactly 1 tail.	$\{HHT, HTH, THH\}$	$3/8$
Toss exactly 1 head.	$\{HTT, THT, TTH\}$	$3/8$
First toss is heads.	$\{HHH, HHT, HTH, HTT\}$	$4/8 = 1/2$
Toss same number of heads as tails.	$\{\ \}$ (*Note*: This is the *impossible event*.)	0
Toss 3 heads or fewer.	S (*Note*: This event is called the *certain event*.)	1

EXAMPLE 15.20 Rolling a Pair of Honest Dice

This example is a follow-up of Examples 15.4 and 15.5. Imagine that you are playing some game that involves rolling a pair of honest dice, and the only thing that matters is the total of the two numbers rolled. As we saw in Example 15.4 the sample space in this situation is $S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$, where the outcomes are the possible totals that could be rolled. This sample space has $N = 11$ possible outcomes, but the outcomes are not equally likely, so it would be wrong to assign to each outcome the probability 1/11. In fact, most people with some experience rolling a pair of dice know that the likelihood of rolling a 7 is much higher than that of rolling a 12.

How can we find the exact probabilities for the various totals that one might roll? We can answer this question by considering the sample space discussed in Example 15.5, where every one of the $N = 36$ possible rolls is listed as a separate outcome (see Fig. 15-1 in Example 15.5). Because the dice are honest, each of these 36 possible outcomes is equally likely to occur, so the probability of each is 1/36. Now we have something!

Table 15-5 shows the probability of rolling a 2, 3, 4, . . . , 12. In each case the numerator represents the number of ways that particular total can be rolled. For example, the event "roll a 7" consists of six distinct possible rolls: {⚀⚅, ⚁⚄, ⚂⚃, ⚃⚂, ⚄⚁, ⚅⚀}. Thus, the probability of "rolling a 7" is $6/36 = 1/6$. The other probabilities are computed in a similar manner.

TABLE 15-5

Event	Probability
"Roll a 2": {⚀⚀}	1/36
"Roll a 3": {⚀⚁, ⚁⚀}	2/36
"Roll a 4": {⚀⚂, ⚁⚁, ⚂⚀}	3/36
"Roll a 5": {⚀⚃, ⚁⚂, ⚂⚁, ⚃⚀}	4/36
"Roll a 6": {⚀⚄, ⚁⚃, ⚂⚂, ⚃⚁, ⚄⚀}	5/36
"Roll a 7": {⚀⚅, ⚁⚄, ⚂⚃, ⚃⚂, ⚄⚁, ⚅⚀}	6/36
"Roll an 8": {⚁⚅, ⚂⚄, ⚃⚃, ⚄⚂, ⚅⚁}	5/36
"Roll a 9": {⚂⚅, ⚃⚄, ⚄⚃, ⚅⚂}	4/36
"Roll a 10": {⚃⚅, ⚄⚄, ⚅⚃}	3/36
"Roll an 11": {⚄⚅, ⚅⚄}	2/36
"Roll a 12": {⚅⚅}	1/36

EXAMPLE 15.21 Rolling a Pair of Honest Dice: Part 2

Once again, we are rolling a pair of honest dice. We are now going to find the probability of the event E: "at least one of the dice comes up an ace." (In dice jargon, a ⚀ is called an ace.)

This is a slightly more sophisticated counting question than the ones in Example 15.20. We will show three different ways to answer the question.

- **Tallying.** We can just write down all the individual outcomes in the event E and tally their number. This approach gives

  $E =$ 

  and $\Pr(E) = 11/36$. There is not much to it, but in a larger sample space it will take a lot of time and effort to list every individual outcome in the event.
- **Complementary Event.** Behind this approach is the germ of a really important idea. Imagine that you are playing a game, and you win if at least one of the two numbers comes up an ace (that's event E). Otherwise you lose (call that event F). The two events E and F are called **complementary events**, and E is called the **complement** of F (and vice versa). The key idea is that *the probabilities of complementary events add up to 1*. Thus, when you want $\Pr(E)$ but it's actually easier to find $\Pr(F)$, then you first do that. Once you have $\Pr(F)$, you are in business: $\Pr(E) = 1 - \Pr(F)$.

  Here we can find $\Pr(F)$ by a direct application of the multiplication principle. There are 5 possibilities for the first die (it can't be an ace but it can be anything else) and likewise there are five possibilities for the second die. This means that there are 25 different outcomes in the event F, and thus $\Pr(F) = 25/36$. It follows that $\Pr(E) = 1 - (25/36) = 11/36$.
- **Independent Events.** In this approach, we look at each die separately and, as usual, pretend that one die is white and the other one is red. We will let F_1 denote the event "the white die does *not* come up an ace" and F_2 denote the event "the red die does *not* come up an ace." Clearly, $\Pr(F_1) = 5/6$ and $\Pr(F_2) = 5/6$. Now comes a critical observation: The probability that both events F_1 and F_2 happen can be found by *multiplying* their respective probabilities. This means that $\Pr(F) = (5/6) \times (5/6) = 25/36$. We can now find $\Pr(E)$ exactly as before: $\Pr(E) = 1 - \Pr(F) = 11/36$.

Of the three approaches used in Example 15.20, the last approach appears to be the most convoluted, but, in fact, it is the one with the most promise. It is based on the concept of *independence* of events. Two events are said to be **independent events** if the occurrence of one event does not affect the probability of the occurrence of the other. When events E and F are independent, the probability that both occur is the product of their respective probabilities; in other words, $\Pr(E \text{ and } F) = \Pr(E) \cdot \Pr(F)$. This is called the **multiplication principle for independent events**.

The multiplication principle for independent events in an important and useful rule, but be forewarned—it works *only* with independent events! For events that are *not* independent, multiplying their respective probabilities gives us a bunch of nonsense.

Our next example illustrates the real power of the multiplication principle for independent events. As we just mentioned, this principle can only be applied when the events in question are independent, and in many circumstances this is not the case. Fortunately, in the examples we are going to consider the independence of the events in question is intuitively obvious. For example, if we roll an honest die several times, what happens on each roll has no impact whatsoever on what is going to happen on subsequent rolls (i.e., the rolls are independent events).

EXAMPLE 15.22 **Rolling a Pair of Honest Dice: Part 3**

Imagine a game in which you roll an honest die four times. If at least one of your rolls comes up an ace, you are a winner. Let E denote the event "you win" and F denote the event "you lose." We will find $\Pr(E)$ by first finding $\Pr(F)$, using the same ideas we developed in Example 15.21.

We will let F_1, F_2, F_3, and F_4 denote the events "first roll is not an ace," "second roll is not an ace," "third roll is not an ace," and "fourth roll is not an ace," respectively. Then

$$\Pr(F_1) = 5/6, \quad \Pr(F_2) = 5/6, \quad \Pr(F_3) = 5/6, \quad \Pr(F_4) = 5/6$$

Now we use the multiplication principle for independent events:

$$\Pr(F) = (5/6) \times (5/6) \times (5/6) \times (5/6) = (5/6)^4 \approx 0.482$$

Finally, we find $\Pr(E)$:

$$\Pr(E) = 1 - \Pr(F) \approx 0.518$$

EXAMPLE 15.23 **Five-Card Poker Hands: Part 2**

One of the best hands in five-card draw poker is a *four-of-a-kind* (four cards of the same value plus a fifth card called a *kicker*). Among all four-of-a-kind hands, the very best is the one consisting of four aces plus a kicker. We will let F denote the event of drawing a hand with four aces plus a kicker in five-card draw poker. (The F stands for "fabulously lucky.") Our goal in this example is to find $\Pr(F)$.

The size of our sample space is $N = {}_{52}C_5 = 2{,}598{,}960$ (see Example 15.13). Of these roughly 2.6 million possible hands, there are only 48 hands in F: four of the five cards are the aces; the kicker can be any one of the other 48 cards in the deck. Thus, $\Pr(F) = 48/2{,}598{,}960 \approx 0.0000185$, roughly 1 in 50,000.

EXAMPLE 15.24 **The Sucker's Bet**

Imagine that a friend offers you the following bet: Toss an honest coin 10 times. If the tosses come out in an even split (5 *Heads* and 5 *Tails*), you win and your friend buys the pizza; otherwise you lose (and you buy the pizza). Sounds like a reasonable offer—let's check it out.

■ You may want to think about this probability and take a guess before you read on — you might be surprised when you see the answer.

We are going to let E denote the event "5 *Heads* and 5 *Tails* are tossed," and we will compute $\Pr(E)$. We already found (Example 15.7) that there are $N = 1024$ equally likely outcomes when a coin is tossed 10 times and that each of these outcomes can be described by a string of 10 Hs and Ts.

We now need to figure out how many of the 1024 strings of Hs and Ts have 5 Hs and 5 Ts. There are of course, quite a few: *HHHHHTTTTT*, *HTHTHTHTHT*, *TTHHHTTTHH*, and so on. Trying to make a list of all of them is not a very practical idea. The trick is to think about these strings as 10 slots each taken up by an H or a T, but once you determine which slots have Hs you automatically know which slots have Ts. Thus, each string of 5 Hs and 5 Ts is determined by the choice of the 5 slots for the Hs. These are unordered choices of 5 out of 10 slots, and the number of such choices is ${}_{10}C_5 = 252$.

Now we are in business: $\Pr(E) = 252/1024 \approx 0.246$.

The moral of this example is that your friend is a shark. When you toss an honest coin 10 times, about 25% of the time you are going to end up with an even split, and the other 75% of the time the split is uneven.

CONCLUSION

While the average citizen thinks of probabilities, chances, and odds as vague, informal concepts that are useful primarily when discussing the weather or playing the lottery, scientists and mathematicians think of probability as a formal framework within which the laws that govern chance events can be understood. The basic elements of this framework are a *sample space* (which represents a precise mathematical description of all the possible outcomes of a *random experiment*), *events* (collections of these outcomes), and a *probability assignment* (which associates a numerical value to each of these events).

Of the many ways in which probabilities can be assigned, a particularly important case is the one in which all individual outcomes have the same probability (*equiprobable spaces*). When this happens, the critical steps in calculating probabilities revolve around two basic (but not necessarily easy) questions: (1) What is the size of the sample space, and (2) what is the size of the event in question? To answer these kinds of questions, knowing the basic principles of "counting" is critical.

When we stop to think how much of our lives is ruled by fate and chance, the importance of probability theory in almost every walk of life is hardly surprising. Understanding the basic mathematical principles behind this theory can help us better judge when taking a chance is a smart move and when it is not. In the long run, this will make us not just better card players, but also better and more successful citizens.

> "His Sacred Majesty, Chance, decides everything."
>
> —Voltaire

PROFILE: Persi Diaconis (1945–)

(a)


(b)


(c)


(d)


(e)


(f)


(g)

(h)

Persi Diaconis picks up an ordinary deck of cards, fresh from the box, and writes a word in Magic Marker on one side: RANDOM. He shuffles the deck once. The letters have re-formed themselves into six bizarre runes that still look vaguely like the letters R, A, and so on. Diaconis shuffles again, and the markings on the side become indecipherable. After two more shuffles, you can't even tell that there used to be six letters. The side of the pack looks just like the static on a television set. It didn't look random before, but it sure looks random now.

Keep watching. After two more shuffles, the word RANDOM miraculously reappears on the side of the deck—only it is written twice, in letters half the original size. After one more shuffle, the original letters materialize at the original size. Diaconis turns the cards over and spreads them out with a magician's flourish, and there they are in their exact original sequence, from the ace of spades to the king of diamonds.

Diaconis has just performed eight perfect shuffles in a row. There's no hocus-pocus, just skill perfected in his youth: Diaconis ran away from home at 14 to become a magician's assistant and later became a professional magician and blackjack player. Even now at 57, he is one of a couple of dozen people on the planet who can do eight perfect shuffles in less than a minute.

Diaconis's work these days involves much more than nimbleness of hand. He is a professor of mathematics and statistics at Stanford University. But he is also the world's leading expert on shuffling. He knows that what seems to be random often isn't, and he has devoted much of his career to exploring the difference. His work has applications to filing systems for computers and the reshuffling of the genome during evolution. And it has led him back to Las Vegas, where, instead of trying to beat the casinos, he now works for them.

A card counter in blackjack memorizes the cards that have already been played to get better odds by making bets based on his knowledge of what has yet to turn up. If the deck has a lot of face cards and 10's left in it, for instance, and he needs a 10 for a good hand, he will bet more because he's more likely to get it. A good card counter, Diaconis estimates, has a 1 to 2 percent advantage over the casino. On a bad day, a good card counter can still lose $10,000 in a hurry. And on a good day, he may get a tap on the shoulder by a large person who will say, "You can call it a day now." By his mid-twenties, Diaconis had figured out that doing mathematics was an easier way to make a living.

Two years ago, Diaconis himself got a tap on the shoulder. A letter arrived from a manufacturer of casino equipment, asking him to figure out whether its card-shuffling machines produced random shuffles. To Diaconis's surprise, the company gave him and his Stanford colleague, Susan Holmes, carte blanche to study the inner workings of the machine. It was like taking a Russian spy on a tour of the CIA and asking him to find the leaks.

When shuffling machines first came out, Diaconis says, they were transparent, so gamblers could actually see the cutting and riffling inside. But gamblers stopped caring after a while, and the shuffling machines turned into closed boxes. They also stopped shuffling cards the way humans do. In the machine that Diaconis and Holmes looked at, each card gets randomly directed, one at a time, to one of 10 shelves. The shuffling machine can put each new card either on the top of the cards already on the shelf or on the bottom, but not between them.

"Already I could see there was something wrong," says Holmes. If you start out with all the red cards at the top of the deck and all the black cards at the bottom, after one pass through the shuffling machine you will find that each shelf contains a red-black sandwich. The red cards, which got placed on the shelves first, form the middle of each sandwich. The black cards, which came later, form the outside. Since there are only 10 shelves, there are at most 20 places where a red card is followed by a black one or vice versa—fewer than the average number of color changes (26) that one would expect from a random shuffle.

The nonrandomness can be seen more vividly if the cards are numbered 1 to 52. After they have passed through the shuffling machine, the numbers on the cards form a zigzag pattern. The top card on the top shelf is usually a high number. Then the numbers decrease until they hit the middle of the first red-black sandwich; then they increase and decrease again, and so on, at most 10 times.

Diaconis and Holmes figured out the exact probability that any given card would end up in any given location after one pass through the machine. But that didn't indicate whether a gambler could use this information to beat the house.

So Holmes worked out a demonstration. It was based on a simple game. You take cards from a deck one by one and each time try to predict what you've selected before you look at it. If you keep track of all the cards, you'll always get the last one right. You'll guess the second-to-last card right half the time, the third-to-last a third of the time, and so on. On average, you will guess about 4.5 cards correctly out of 52.

By exploiting the zigzag pattern in the cards that pass through the shuffling machine, Holmes found a way to double the success rate. She started by predicting that the highest possible card (52) would be on top. If it turned out to be 49, then she predicted 48—the next highest number—for the second card. She kept going this way until her prediction was too low—predicting, say, 15 when the card was actually 18. That meant the shuffling machine had reached the bottom of a zigzag and the numbers would start climbing again. So she would predict 19 for the next card. Over the long run, Holmes (or, more precisely, her computer) could guess nine out of every 52 cards correctly.

To a gambler, the implications are staggering. Imagine playing blackjack and knowing one-sixth of the cards before they are turned over! In reality, a blackjack player would not have such a big advantage, because some cards are hidden and six full decks are used. Still, Diaconis says, "I'm sure it would double or triple the advantage of the ordinary card counter."

Diaconis and Holmes offered the equipment manufacturer some advice: Feed the cards through the machine twice. The alternative would be more expensive: Build a 52-shelf machine.

A small victory for shuffling theory, one might say. But randomization applies to more than just cards. Evolution randomizes the order of genes on a chromosome in several ways. One of the most common mutations is called a "chromosome inversion," in which the arm of a chromosome gets cut in two random places, flipped over end-to-end, and reattached, with the genes in reverse order. In fruit flies, inversions happen at a rate of roughly one per every million years. This is very similar to a shuffling method called transposition that Diaconis studied 20 years ago. Using his methods, mathematical biologists have estimated how many inversions it takes to get from one species of fruit fly to another, or to a completely random genome. That, Diaconis suggests, is the real magic he ran away from home to find. "I find it amazing," he says, "that mathematics developed for purely aesthetic reasons would mesh perfectly with what engineers or chromosomes do when they want to make a mess."

Source: Reprinted by permission from Dana Mackenzie, "The Mathematics of Shuffling," *Discover* (October 2002), pp. 22–23.

KEY CONCEPTS

combinations, **564**
complement, **571**
complementary event, **571**
equiprobable space, **569**
event, **566**
impossible event, **566**
independent events, **571**
multiplication principle for independent events, **571**
multiplication rule, **561**
odds against, **573**
odds of (odds in favor of), **573**
permutations, **564**
probability assignment, **567**
probability space, **568**
random experiment, **558**
sample space, **558**
simple event, **566**