# 16 Probabilities, Odds, and Expectations

## Measuring Uncertainty and Risk

Friday, March 30, 2012, was a really bad day for worker productivity—even worse than the Fridays before Christmas. In many parts of the country people started lining up at the crack of dawn outside liquor stores and convenience stores and waited for hours for a chance to get inside. What for? To buy a dream.

That day the Mega Millions Lottery jackpot hit $640 million—the largest jackpot ever recorded in the United States—and lottery fever swept the nation. Buy a $1 lottery ticket and make $640 million? Who could argue with that deal? Waiting in line for hours? A small price to pay.

“His sacre
Chance,
everythin

The natural impulse is to count the number of choices using the multiplication rule (and a box model) as shown in Fig. 16-5. This would give an answer of $31 \times 30 = 930$ true doubles. Unfortunately, this answer is *double counting* each of the true doubles. Why?

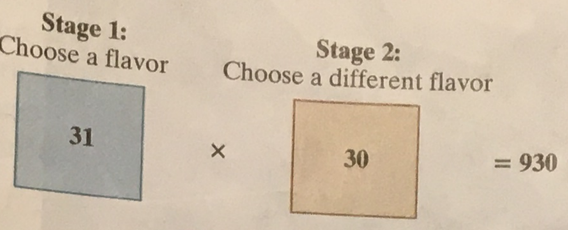



FIGURE 16-5

When we use the multiplication rule, *there is a well-defined order to things*, and a scoop of vanilla followed by a scoop of chocolate is counted separately from a scoop of chocolate followed by a scoop of vanilla. But by all reasonable standards, a bowl with a scoop of vanilla plus a scoop of chocolate is the same regardless of which scoop goes in the bowl first. (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 possible true doubles is $(31 \times 30)/2 = 465$.

Example 16.13 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 16.14** **ICE CREAM: EPISODE 4**

Some days you have a real craving for ice cream, and on such days you like to go to the ice cream parlor around the corner and order a "true triple." A "true triple" consists of three scoops of ice cream each of a different flavor. How many different "true triples" are there?

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 $26{,}970/6 = 4495$.

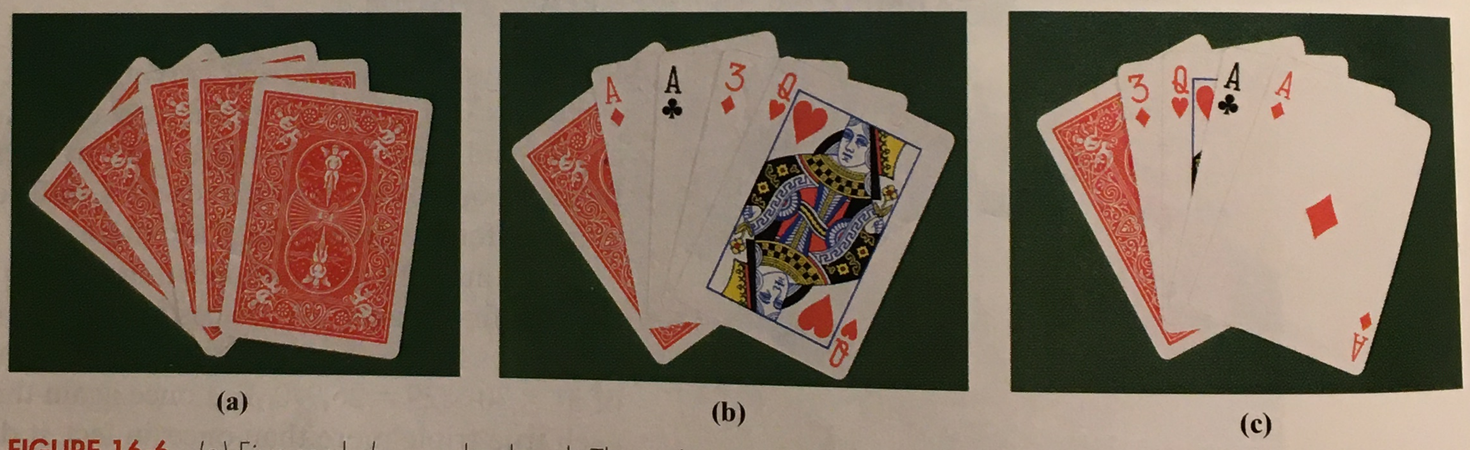
Why is it that in the total of 26,970 we are counting each true triple *six* times? The answer comes down to this: Any three flavors (call them $X$, $Y$, and $Z$) can be listed in 6 different ways ($XYZ, XZY, YXZ, YZX, ZXY$, and $ZYX$). The multiplication rule counts each of these separately, but when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

The bottom line is that with 31 different flavors we can create 4495 different true triples. 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 16.15 FIVE-CARD POKER HANDS: EPISODE 1

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* [Fig. 16-6(a)] the order in which the cards are dealt is irrelevant; in five-card *stud* [Figs. 16-6(b) and (c)] 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.



(a) (b) (c)

**FIGURE 16-6** (a) Five-card *draw* poker hand: The order in which the cards are dealt is irrelevant. (b) and (c) Five-card *stud* poker hands: same up-cards but dealt in a different sequence.

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 16.14 and 16.15. Suppose that we have a set of $n$ distinct objects and we want to choose $r$ different objects from this set. The number of ways that this can be done depends on whether the choices are *ordered* or *unordered*. Ordered choices are the generalization of stud poker hands—when the same objects are chosen in a different order you get a different outcome. Unordered choices are the generalization of draw poker hands—the order in which the objects are chosen does not affect the outcome. To distinguish between these two scenarios, we use the terms **permutation** to describe an ordered selection and **combination** to describe an unordered selection.

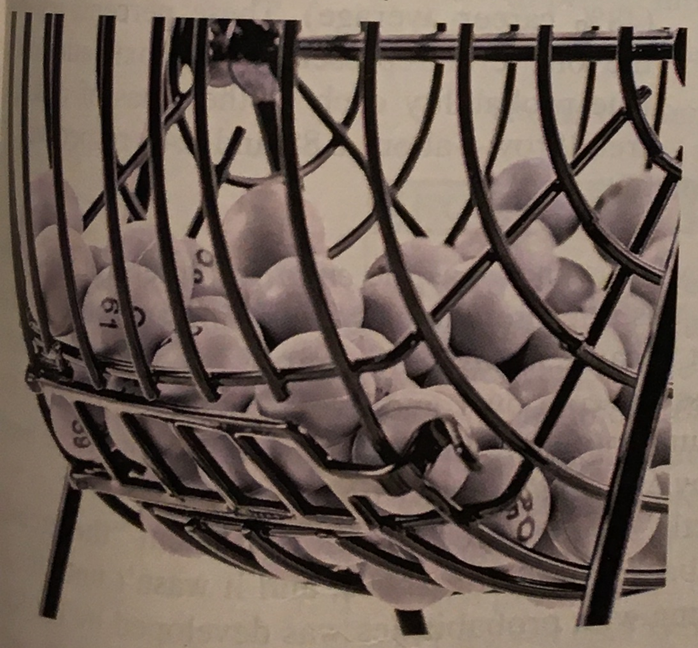
- **Permutation.** An *ordered* selection of $r$ objects chosen from a set of $n$ objects ($r$ can be any integer between 0 and $n$). In the special case where $r = n$ all the objects are selected and the permutation is just a rearrangement or shuffing of the objects.
- **Combination.** An *unordered* selection of $r$ objects chosen from a set of $n$ objects ($r$ can be any integer between 0 and $n$).

We will use the standard notation ${}_nP_r$ to denote the *number of permutations* of $r$ objects chosen from a set of $n$ objects and ${}_nC_r$ to denote the number of *combinations* of $r$ objects chosen from a set of $n$ objects. (Some calculators use variations of this notation, such as $P_{n,r}$ and $C_{n,r}$, respectively.)

The numbers ${}_nP_r$ and ${}_nC_r$ are typically found using a scientific calculator, but it is always a good idea to know the basic formulae behind the computations:

- ${}_nP_r = n(n-1)(n-2)\dots(n-r+1)$, or equivalently, ${}_nP_r = \frac{n!}{(n-r)!}$
- ${}_nC_r = {}_nP_r/r!$, or equivalently, ${}_nC_r = \frac{n!}{r!(n-r)!}$

## EXAMPLE 16.16 THE MEGA MILLIONS LOTTERY: EPISODE 1



A typical lottery drawing consists of taking a few balls out of a larger set of numbered balls inside a spinning container called a *hopper*. The range of numbers on the balls and the number of balls drawn from the hopper can vary with various kinds of lotteries. If the numbers drawn from the hopper match the numbers on your lottery ticket, you win the jackpot. You can also win smaller prizes if some (but not all) of your numbers match some of the numbers drawn.

For the Mega Millions lottery, the drawing comes in two parts. First, five balls are drawn from a hopper containing white balls numbered 1 through 56. Then, one ball—called the "mega ball"—is drawn from a different hopper containing gold balls numbered 1 through 46. In this example we will accurately count all possible Mega Millions drawings, even though this is a very large number.

We start counting the number of drawings of five balls from the first hopper, containing balls numbered 1 through 56. Since the order in which the numbered balls come out of the hopper is irrelevant (the only things that matters is if you have a match or not), we are dealing with combinations of five numbers taken from the set of numbers 1 through 56. The number of such combinations is ${}_{56}C_5 = 3{,}819{,}816$.

As for the mega ball, there are 46 balls in the second hopper and only one comes out, so there are 46 possibilities. Using the multiplication rule, we can conclude that the number of possible Mega Millions drawings is $({}_{56}C_5) \times 46 = 175{,}711{,}536$.

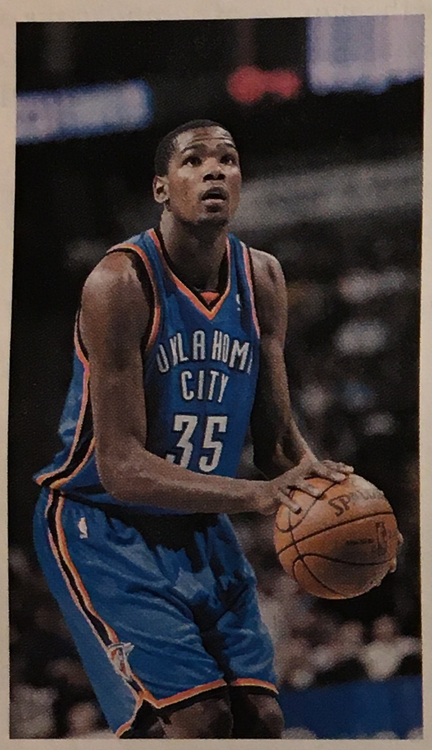
# 16.3 Probabilities and Odds

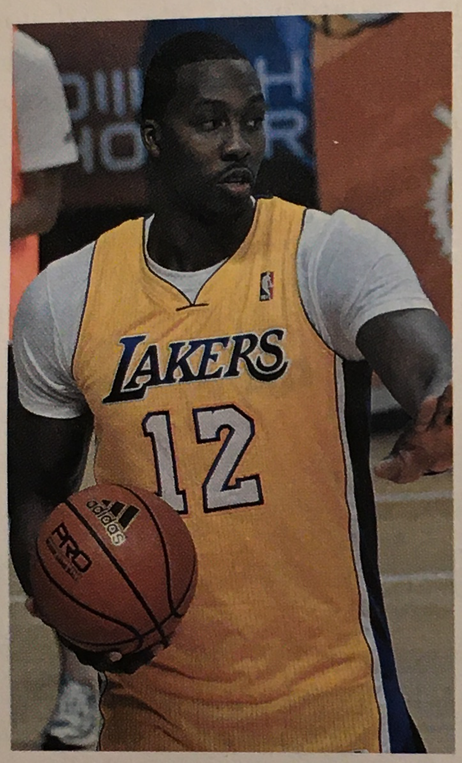
If we toss a coin in the air, *what is the probability that it will land heads up?* This one is not a very profound question, and almost everybody agrees on the answer, although not necessarily for the same reason. The standard answer given is 1 out of 2, or 1/2. But why is the answer 1/2, and what does 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 16.17 SHOOTING FREE THROWS: EPISODE 2




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 (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 that player—it makes a difference whether the shooter is a Kevin Durant or a Dwight Howard. Durant is one of the best free-throw shooters in the NBA, with a career average of 88%, while Howard is a notoriously poor free-throw shooter (48% 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 0.88 and 0.48, respectively. In either case, the probability is not 0.5.

Example 16.17 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 this section we will introduce some of the basic ideas behind the mathematical theory of probability.

### Probability Assignments

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

| Probability |
|---|
| 0 |
| $p$ |
| $1-p$ |
| 1 |

16-3

**EXAMPLE 16.18** **SHOOTING FREE THROWS: EPISODE 3**

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 Kevin Durant, Dwight Howard, or Joe Schmoe.

How can we describe the probability that this unknown free-throw shooter 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 16-3 shows the probability of each possible event in the sample space $S = \{s, f\}$. Basically, Table 16-3 gives a generic model for free-throw shooting. It works when the free-throw shooter is Kevin Durant, Dwight Howard, or Joe Schmoe. Each one of the choices results in a different assignment of numbers to the events in the sample space.

Example 16.18 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:

- All probabilities are numbers between 0 and 1
- The sum of the probabilities of the simple events equals 1.

The next example illustrates how a probability assignment works.

**EXAMPLE 16.19** **HANDICAPPING A TENNIS TOURNAMENT: EPISODE 1**

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 16.19 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 summarizes the key 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 $\Pr(S) = 1$.]

## 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 $k/N$.

- **Probability of an event:** If $k$ denotes the size of an event $E$ and $N$ denotes the size of the sample space $S$, then in an equiprobable space $\Pr(E) = k/N$.

### EXAMPLE 16.20 ROLLING A PAIR OF DICE: EPISODE 2

This example is a follow-up of Example 16.4. Imagine that you are playing a 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 16.4, 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.

To determine the probability of each total, we consider the roll of each die separately and look at the sample space consisting of the 36 possible ways in which one

can roll a pair of dice (one white and one red) as discussed in Example 16.4 (see Fig. 16-1, page 482). Because the dice are honest, each of these 36 possible outcomes is equally likely to occur, so the probability of each is 1/36.

Table 16-4 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.

| 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 |

■ TABLE 16-4

**EXAMPLE 16.21** ROLLING A PAIR OF DICE: EPISODE 3

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 16.20. 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.** 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 16.21, 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, $\Pr(E \text{ and } F) = \Pr(E)\Pr(F)$. 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 16.22 ROLLING MORE DICE

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 16.21.

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 16.23** **FIVE-CARD POKER HANDS: EPISODE 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 16.15). 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 16.24** **A CAR AND TWO GOATS: EPISODE 2**

This example is a follow-up of Example 16.5. You are a contestant in a television game show and you have to choose one of three doors to win either a fancy sports car (behind one of the doors) or a goat (behind each of the other two). For starters, the sample space for the game consists of nine outcomes: $S = \{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3)\}$, where $(x, y)$ denotes that you chose Door #$x$, and the car is behind Door #$y$. Assuming that the car is randomly placed behind one of the three doors and that you make a random choice of door, each of the nine outcomes has an equal probability of 1/9. The set of outcomes $E = \{(1,1), (2,2), (3,3)\}$ represents the event $E =$ "you win the car"; the remaining set of outcomes $F = \{(1,2), (1,3), (2,1), (2,3), (3,1), (3,2)\}$ represents the complementary event $F =$ "you win a goat." This formalizes what is everyone's intuition: Your probability of winning a car is $\Pr(E) = 1/3$; your probability of winning a goat is $\Pr(F) = 2/3$.

So let's say that you choose Door #1. At this point nothing has changed, so your probability of winning a car is still 1/3. Now comes a new wrinkle: before showing what is behind Door #1 the host—who knows where the car is—opens Door #3 to show there is a goat behind it, and offers you the option of switching. Should you *stick* with your original choice, or *switch* to Door #2? Most people's intuition is that it makes no difference—stick or switch, your probability of winning the car is the same. This is a notorious example where people's intuitions turn out to be wrong.

If you stick with Door #1 then you are essentially disregarding the information provided by the host (that the car is not behind Door #3). Your probability of winning the car is still 1/3, and your probability of winning a goat is still 2/3.

Now let's find the probabilities if you switch to Door #2. The key here is to consider the reasons why the host opened Door #3. Obviously the host would never open the door showing the car, so if the car happens to be behind Door #2, the host is forced to open Door #3. In this case, the host opens Door #3 with probability 1 (2 out of 2 times). On the other hand, if the car is behind Door #1 the host could have chosen to open either Door #3 or Door #2. Assuming that he chooses each door with probability 1/2, when the car is behind Door #1 he opens Door #3 half of the time (1 out of 2 times). This means that once the host opened Door #3, 2 out of every 3 times he does it because he has no choice (as the car is behind Door #2), and 1 out of 3 times he does it because he randomly chooses it (and the car is behind Door #1). It follows that when you switch, the probability of winning the car is 2/3; the probability of winning a goat is 1/3. You should definitely switch!

## Odds

Dealing with probabilities as numbers that are always between 0 and 1 is the mathematician's way of having a consistent terminology. To the everyday user, consistency is not that much of a concern, and we know that people often express the likelihood of an event in terms of *odds*. Probabilities and odds are not identical concepts but are closely related.

- **Odds.** Let $E$ be an arbitrary event. If $F$ denotes the number of ways that event $E$ can occur (the *favorable* outcomes or *hits*) and $U$ denotes the number of ways that event $E$ does not occur (the *unfavorable* outcomes, or *misses*), then the **odds of** (also called the **odds in favor of**) the event $E$ are given by the ratio $F$ to $U$, and the **odds against** the event $E$ are given by the ratio $U$ to $F$. When $F$ and $U$ have a common factor, it is customary (and desirable) to simplify the numbers before giving the odds. One rarely hears odds of 15 to 5 — they are much better expressed as odds of 3 to 1.

### EXAMPLE 16.25 ODDS OF MAKING FREE THROWS

In Example 16.17 we discussed the fact that Kevin Durant shoots free throws with a probability of $p = 0.88$. We can interpret this to mean that on the average, out of every 100 free throws, Durant is going to make 88 and miss 12, for a hit/miss ratio of 88 to 12, or 22 to 3. These are the odds of Durant making a free throw. We can also express the odds in a negative context and say that the *odds against* Durant making the free throw are 3 to 22.

In contrast, the odds of Dwight Howard making a free throw can be described as being *roughly* 1 to 1. To be more precise, Howard shoots free throws with a probability of $p = 0.48$ (see Example 16.17), which represents odds of 12 to 13.

### EXAMPLE 16.26 ODDS OF ROLLING A "NATURAL"

Suppose that you are playing a game in which you roll a pair of dice, presumably honest. In this game, when you roll a "natural" (i.e., roll a 7 or an 11) you automatically win.

If we let $E$ denote the event "roll a natural," we can check that out of 36 possible outcomes 8 are favorable (6 ways to "roll a 7" and two ways to "roll an 11"—see Table 16-4) and the other 28 are unfavorable. It follows that the odds of rolling a "natural" are 2 to 7 (originally 8 to 28 and then simplified to 2 to 7).

We can convert odds into probabilities, and probabilities into odds using the following two rules:

- If the odds of the event $E$ are $F$ to $U$, then $\Pr(E) = F/(F+U)$.
- If $\Pr(E) = A/B$, then the odds of $E$ are $A$ to $B - A$. (When the probability is given in decimal form, the best thing to do is to first convert the decimal form into fractional form.)

### EXAMPLE 16.27 HANDICAPPING A TENNIS TOURNAMENT: EPISODE 2

This example is a follow-up to Example 16.19. Recall that the probability assignment for the tennis tournament was as follows: $\Pr(A) = 0.08$, $\Pr(B) = 0.16$, $\Pr(C) = 0.20$, $\Pr(D) = 0.25$, $\Pr(E) = 0.16$, and $\Pr(F) = 0.15$. We will now express each of these probabilities as odds. (Notice that to do so we first convert the decimals into fractions in reduced form.)

- $\Pr(A) = 0.08 = 8/100 = 2/25$. Thus, the odds of $A$ winning the tournament are 2 to 23.
- $\Pr(B) = 0.16 = 16/100 = 4/25$. The odds of $B$ winning the tournament are 4 to 21.
- $\Pr(C) = 0.20 = 20/100 = 1/5$. The odds of $C$ winning the tournament are 1 to 4.
- $\Pr(D) = 0.25 = 25/100 = 1/4$. The odds of $D$ winning the tournament are 1 to 3.
- $\Pr(E) = \Pr(B) = 4/25$. The odds of E winning the tournament are the same as $B$'s (4 to 21).
- $\Pr(F) = 0.15 = 15/100 = 3/20$. The odds of $F$ winning the tournament are 3 to 17.

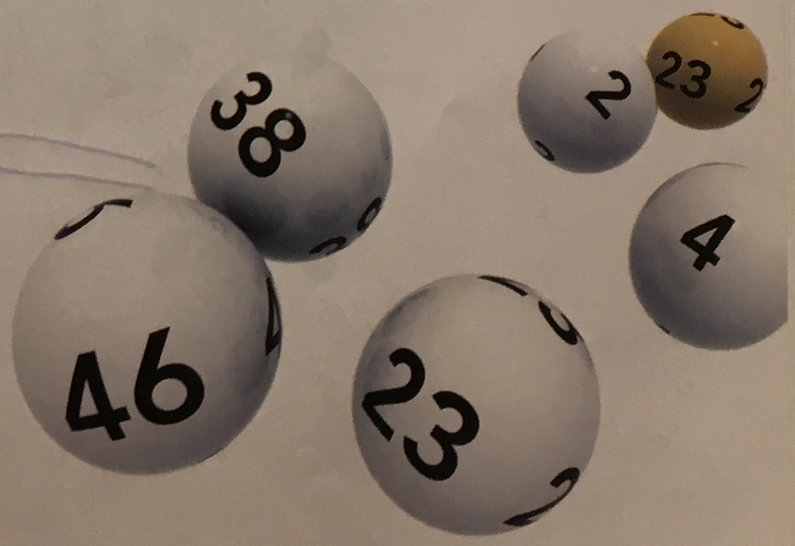
His sacred majesty,
Chance, decides
everything. ”

– *Voltaire*

The March 30, 2012, lottery frenzy created by the Mega Millions jackpot (by the way, the final jackpot of $656 million was split between three different winners) is an extreme example of the combination that makes big prize lotteries so appealing all over the world: low risk–high reward. Spend $1 for a lottery ticket, make millions if you win. But there is a key missing element in this equation—what are the chances (probabilities, odds) that you *will win*? Without taking into account this third variable, the low risk–high reward lottery formula is just fool's gold.

By the time we are finished with this chapter we will be able to understand how risk, rewards, and probabilities are combined in a precise mathematical way that allows us to answer major existential questions like, "Should I really spend my money buying lottery tickets?" To get there, we will have to cover a lot of ground quickly—by necessity, this chapter covers a lot of important ideas in just a few pages. The chapter starts with an introduction to *sample spaces* and *events*—the basic building blocks for computing probabilities (Section 16.1). Section 16.2 is a quick introduction (for many readers a review) of the *multiplication rule*, *permutations*, and *combinations*—these are key mathematical tools for probability calculations. In Section 16.3 we finally get to *probabilities* and *odds* (they are two different formulations of the same idea) with plenty of examples that illustrate how to calculate both. In Section 16.4 we introduce the concept of *expectation* (or *expected value*), the key concept in understanding the relationship between risk and reward. In Section 16.5 we use several real-world examples (including lotteries) to illustrate how to use expectation as a tool for measuring risk.

*Disclaimer:* Throughout the chapter we will use many examples of games of chance—craps, poker, horse races, lotteries, etc. Games of chance often serve as the ideal setting in which to analyze and deal with interesting probability questions. This should not be taken in any way, shape, or form as an encouragement to gamble or an endorsement of gambling—on the contrary, several of the examples in the chapter are there to illustrate mathematically why gambling, in general, is not such a good idea.



*A final word of caution:* There is a difference between odds as discussed in this section and the *payoff odds* posted by casinos or bookmakers in sports gambling situations. Suppose we read in the newspaper, for example, that the Las Vegas bookmakers have established that "the odds that the Boston Celtics will win the NBA championship are 5 to 2." What this means is that if you want to bet in favor of the Celtics, for every $2 that you bet, you can win $5 if the Celtics win. This ratio may be taken as some indication of the actual odds in favor of the Celtics winning, but several other factors affect payoff odds, and the connection between payoff odds and actual odds is tenuous at best.

# 16.4 Expectations

In this section we will use probabilities to analyze mathematically the measurement of risks. We start with an introduction to the important concept of a weighted average.

## Weighted Averages

**EXAMPLE 16.28** **COMPUTING CLASS SCORES**

Imagine that you are a student in Prof. Goodie's Stat 100 class. The grading for the class is based on two midterms, homework, and a final exam. The breakdown for the scoring is given in the first two rows of Table 16-5. Your scores are given in the last row. You needed to average 90% or above to get an A in the course. Did you?

| | Midterm 1 | Midterm 2 | Homework | Final exam |
|---|---|---|---|---|
| Weight | 20% | 20% | 25% | 35% |
| Possible points | 100 | 100 | 200 | 200 |
| Your scores | 82 | 88 | 182 | 190 |

■ TABLE 16-5

One of your friends says you should add all your scores and divide by the total number of possible points to compute your final average. This gives $(82 + 88 + 182 + 190)/600 = 0.9033\cdots \approx 90.33\%$.

If your friend is right you would end up with an A. A second friend says that you are going to miss an A by just one percentage point, and isn't that too bad! Your second friend's calculation is as follows: You got 82% on the first midterm, 88% on the second midterm, 91% (182/200) on the homework, and 95% (190/200) on the final exam. The average of these four percentages is $(82\% + 88\% + 95\% + 91\%)/4 = 89\%$.

Fortunately, you know more math than either one of your friends. The correct computation, you explain to them patiently, requires that we take into account that (1) the scores are based on different scales (100 points, 200 points) and (2) the weights of the scores (20%, 20%, 25%, 35%) are not all the same. To take care of (1) we express the numerical scores as percentages (82%, 88%, 91%, 95%), and to take care of (2) we multiply these percentages by their respective weights ($20\% = 0.20$, $25\% = 0.25$, $35\% = 0.35$). We then add all of these numbers. Thus, your correct average is

$$0.2 \times 82\% + 0.2 \times 88\% + 0.25 \times 91\% + 0.35 \times 95\% = 90\%$$

How sweet it is—you are getting that A after all!

## EXAMPLE 16.29 AVERAGE GPAS AT TASMANIA STATE UNIVERSITY

Table 16-6 shows GPAs at Tasmania State University broken down by class. Our task is to compute the overall GPA of *all undergraduates* at TSU.

| Class | Freshman | Sophomore | Junior | Senior |
|---|---|---|---|---|
| Average GPA | 2.75 | 3.08 | 2.94 | 3.15 |

■ **TABLE 16-6** GPAs at Tasmania State University by class

The information given in Table 16-6 is not enough to compute the overall GPA because the number of students in each class is not the same. After a little digging we find out that the 15,000 undergraduates at Tasmania State are divided by class as follows: 4800 freshmen, 4200 sophomores, 3300 juniors, and 2700 seniors.

To compute the overall school GPA we need to "weigh" the average GPA for each class with the relative size of that class. The relative size of the freshman class is $4800/15{,}000 = 0.32$, the relative size of the sophomore class is $4200/15{,}000 = 0.28$, and so on (0.22 for the junior class and 0.18 for the senior class). When all is said and done, the overall school GPA is

$$0.32 \times 2.75 + 0.28 \times 3.08 + 0.22 \times 2.94 + 0.18 \times 3.15 = 2.9562$$

Examples 16.28 and 16.29 illustrate how to average numbers that have different relative *weights* using the notion of a *weighted average*. In Example 16.28 the scores (82%, 88%, 91%, and 95%) were multiplied by their corresponding "weights" (0.20, 0.20, 0.25, and 0.35) and then added to compute the *weighted average*. In Example 16.29 the class GPAs (2.75, 3.08, 2.94, and 3.15) were multiplied by the corresponding class "weight" (0.32, 0.28, 0.22, 0.18) to compute the *weighted average*. Note that in both examples the *weights* add up to 1, which is not surprising since they represent percentages that must add up to 100%.

■ **Weighted average:** Let $X$ be a variable that takes the values $x_1, x_2, \ldots, x_N$, and let $w_1, w_2, \ldots, w_N$ denote the respective weights for these values, with $w_1 + w_2 + \cdots + w_N = 1$. The **weighted average** for $X$ is given by

$$w_1 \cdot x_1 + w_2 \cdot x_2 + \cdots + w_N \cdot x_N$$

The idea of a weighted average is particularly useful when the weights represent probabilities.

## EXAMPLE 16.30 GUESSING ANSWERS IN THE SAT: EPISODE 1

In the multiple-choice sections of the SAT each question has five possible answers ($A$, $B$, $C$, $D$, and $E$). A correct answer is worth 1 point, and, to discourage indiscriminate guessing, an incorrect answer carries a penalty of 1/4 point (i.e., it counts as $-1/4$ point).

Imagine you are taking the SAT and are facing a multiple-choice question for which you have no clue as to which of the five choices might be the right answer—they all look equally plausible. You can either play it safe and leave it blank, or you can gamble and take a wild guess. In the latter case, the upside of getting 1 point must be measured against the downside of getting a penalty of 1/4 point. What should you do?

Table 16-7 summarizes the possible options and their respective payoffs (for the purposes of illustration assume that the correct answer is $B$, but of course, you don't know that when you are taking the test).

| Option | Leave blank | A | B | C | D | E |
|---|---|---|---|---|---|---|
| Payoff | 0 | −0.25 | 1 | −0.25 | −0.25 | −0.25 |

■ TABLE 16-7

If you are randomly guessing the answer to this multiple-choice question you are unwittingly conducting a *random experiment* with sample space $S = \{A, B, C, D, E\}$. Since each of the five choices is equally likely to be the correct answer (remember, you are clueless on this one), you assign equal probabilities of $p = 1/5 = 0.2$ to each outcome. This probability assignment, combined with the information in Table 16-7, gives us Table 16-8.

| Outcome | Correct answer ($B$) | Incorrect answer ($A$, $C$, $D$, or $E$) |
|---|---|---|
| Point payoff | 1 | −0.25 |
| Probability | 0.2 | 0.8 |

■ TABLE 16-8

Using Table 16-8 we can compute something analogous to a weighted average for the point payoffs with the *probabilities as weights*. We will call this the *expected payoff*. Here the expected payoff $E$ is

$$E = 0.2 \times 1 + 0.8 \times (-0.25) = 0.2 - 0.2 = 0 \text{ points}$$

The fact that the expected payoff is 0 points implies that this guessing game is a "fair" game—in the long term (if you were to make these kinds of guesses many times) the penalties that you would accrue for wrong guesses are neutralized by the benefits that you would get for your lucky guesses. This knowledge will not impact what happens with an individual question [the possible outcomes are still a +1 (lucky) or a −0.25 (wrong guess)], but it gives you some strategically useful information: On the average, totally random guessing on the multiple-choice section of the SAT neither helps nor hurts!

## EXAMPLE 16.31 GUESSING ANSWERS IN THE SAT: EPISODE 2

Example 16.30 illustrated what happens when the multiple-choice question has you completely stumped—each of the five possible choices ($A$, $B$, $C$, $D$, or $E$) looks equally likely to be the right answer. To put it bluntly, you are clueless! At a slightly better place on the ignorance scale is a question for which you can definitely rule out one or two of the possible answers. Under these circumstances we must do a different calculation for the risk/benefits of guessing.

Let's consider first a multiple-choice question for which you can safely rule out one of the five choices. For the purposes of illustration let's assume that the correct answer is $B$ and that you can definitely rule out choice $E$. Among the four possible choices ($A, B, C,$ or $D$) you have no idea which is most likely to be the correct answer, so you are going to randomly guess. In this scenario we assign the same probability (0.25) to each of the four choices, and the guessing game is described in Table 16-9.

| Outcome | Correct answer ($B$) | Incorrect answer ($A$, $C$, or $D$) |
|---|---|---|
| Point payoff | 1 | −0.25 |
| Probability | 0.25 | 0.75 |

■ TABLE 16-9

Once again, the *expected payoff* $E$ can be computed as a weighted average:

$$E = 0.25 \times 1 + 0.75 \times (-0.25) = 0.25 - 0.1875 = 0.0625$$

An expected payoff of 0.0625 points is very small—it takes 16 guesses of this type to generate an expected payoff equivalent to 1 correct answer (1 point). At the same time, the fact that it is a positive expected payoff means that the benefit justifies (barely) the risk.

A much better situation occurs when you can rule out two of the five possible choices (say, $D$ and $E$). Now the random experiment of guessing the answer is described in Table 16-10. [Here it's a little more convenient to express everything in terms of fractions.]

| Outcome | Correct answer ($B$) | Incorrect answer ($A$ or $C$) |
|---|---|---|
| Point payoff | 1 | $-1/4$ |
| Probability | 1/3 | 2/3 |

■ TABLE 16-10

In this situation the expected payoff $E$ is given by

$$E = (1/3) \times 1 + (2/3) \times (-1/4) = 1/3 - 1/6 = 1/6$$

Examples 16.30 and 16.31 illustrate the mathematical reasoning behind a commonly used piece of advice given to SAT-takers: *Guess the answer if you can rule out some of the options; otherwise, don't bother.*

The basic idea illustrated in Examples 16.30 and 16.31 is that of an *expectation* (or *expected value*).

■ **Expectation:** Suppose $X$ is a variable that takes on numerical values $x_1, x_2, x_3, \ldots, x_N$, with probabilities $p_1, p_2, p_3, \ldots, p_N$, respectively. The *expectation* (or *expected value*) of $X$ is given by

$$E = p_1 \cdot x_1 + p_2 \cdot x_2 + p_3 \cdot x_3 + \cdots + p_N \cdot x_N$$

## 16.5 Measuring Risk

In many real-life situations, we face decisions that can have many different potential consequences—some good, some bad, some neutral. These kinds of decisions are often quite hard to make because there are so many intangibles, but sometimes the decision comes down to a simple question: Is the reward worth the risk? If we can quantify the risks and the rewards, then we can compute the *expectation* (expected value) of the outcome to help us make the right decision. The classic illustration of this type of situation is provided by "games" in which there is money riding on the outcome. This includes not only typical gambling situations (dice games, card games, lotteries, etc.) but also buying life insurance or playing the stock market.

Playing the stock market is a legalized form of gambling in which people make investment decisions (buy? sell? when? what?) instead of rolling the dice or drawing cards. Some people have a knack for making good investment decisions and in the long run can do very well—others, just the opposite. A lot has to do with what drives the decision-making process—gut feelings, rumors, and can't-miss tips from your cousin Vinny at one end of the spectrum; reliable information combined with sound mathematical principles at the other end. Which approach would you rather trust your money to? (If your answer is the former, then you might as well skip the next example.)

**EXAMPLE 16.32** **TO BUY OR NOT TO BUY?**

Fibber Pharmaceutical is a new start-up in the pharmaceutical business. Its stock is currently selling for $10. Fibber's future hinges on a new experimental vaccine it believes has great promise for the treatment of the avian flu virus. Before the vaccine can be approved by the Food and Drug Administration (FDA) for use with the general public it must pass a series of clinical trials known as Phase I, Phase II, and Phase III trials. If the vaccine passes all three trials and is approved by the FDA, shares of Fibber are predicted to jump tenfold to $100 a share. At the other end of the spectrum, the vaccine may turn out to be a complete flop and fail Phase I trials. In this case, Fibber's shares will be worthless. In between these two extremes are two other possibilities: The vaccine will pass Phase I trials but fail Phase II trials or will pass Phase I and Phase II trials but fail Phase III trials. In the former case shares of Fibber are expected to drop to $5 a share; in the latter case shares of Fibber are expected to go up to $15 a share.

Table 16-11 summarizes the four possible outcomes of the clinical trials. The last row of Table 16-11 gives the probability of each outcome based on previous experience with similar types of vaccines.

| Outcome of trials | Fail Phase I | Fail Phase II | Fail Phase III | FDA approval |
|---|---|---|---|---|
| Estimated share price | $0 | $5 | $15 | $100 |
| Probability | 0.25 | 0.45 | 0.20 | 0.10 |

■ TABLE 16-11

Combining the second and third rows of Table 16-11, we can compute the expected value $E$ of a future share of Fibber Pharmaceutical:

$$E = 0.25 \times \$0 + 0.45 \times \$5 + 0.20 \times \$15 + 0.10 \times \$100 = \$15.25$$

To better understand the meaning of the $15.25 expected value of a $10 share of Fibber Pharmaceutical, imagine playing the following game: For a cost of $10 you get to roll a die. This is no ordinary die—this die has only four faces (labeled I, II, III, and IV), and the probabilities of each face coming up are different (0.25, 0.45, 0.20, and 0.10, respectively). If you roll a I, your payoff is $0 (your money is gone); if you roll a II, your payoff is $5 (you are still losing $5); if you roll a III, your payoff is $15 (you made $5); and if you roll a IV, your payoff is $100 (jackpot!). The beauty of this random experiment is that it can be played over and over, hundreds or even thousands of times. If you do this, sometimes you'll roll a I and poof—your money is gone, a few times you'll roll a IV and make out like a bandit, other times you'll roll a III or a II and make or lose a little money. The key observation is that if you play this game long enough, the average payoff is going to be $15.25 per roll of the die, giving you an average profit (gain) of $5.25 per roll. In purely mathematical terms, this game is a game well worth playing (but since this book does not condone gambling, this is just a theoretical observation).

Is Fibber Pharmaceutical then a great risk? At first glance, an expected value of $15.25 per $10 invested looks like a great risk, but we have to balance this with the several years that it might take to collect on the investment (unlike the die game that pays off right away). For example, assuming seven years to complete the clinical trials, an investment in Fibber shares has a comparable expected payoff as a seven-year bond with a fixed annual yield of about 6.25%.

### 16.5 Measuring Risk

- **fair (even) game:** a game where no player (including "the house" when there is one) has a built-in advantage over the other players, **503**

## EXERCISES

### WALKING

### 16.1 Sample Spaces and Events

1. Using set notation, write out the sample space for each of the following random experiments.
   (a) A coin is tossed three times in a row. The observation is how the coin lands ($H$ or $T$) on each toss.
   (b) A basketball player shoots three consecutive free throws. The observation is the result of each free throw ($s$ for success, $f$ for failure).
   (c) A coin is tossed three times in a row. The observation is the number of times the coin comes up tails.
   (d) A basketball player shoots three consecutive free throws. The observation is the number of successes.

2. Using set notation, write out the sample space for each of the following random experiments.
   (a) A coin is tossed four times in a row. The observation is how the coin lands ($H$ or $T$) on each toss.
   (b) A student randomly guesses the answers to a four-question true-or-false quiz. The observation is the student's answer ($T$ or $F$) for each question.
   (c) A coin is tossed four times in a row. The observation is the percentage of tosses that are heads.
   (d) A student randomly guesses the answers to a four-question true-or-false quiz. The observation is the percentage of correct answers in the test.

3. Using set notation, write out the sample space for each of the following random experiments:
   (a) Roll a pair of dice. The observation is the difference (in absolute value) of the two numbers that are rolled.
   (b) Toss a coin five times. The observation is the difference (# of heads – # of tails) in the five tosses.

4. Using set notation, write out the sample space for each of the following random experiments:
   (a) Three runners ($A$, $B$, and $C$) are running in a race (assume that there are no ties). The observation is the order in which the three runners cross the finish line.
   (b) Four runners ($A$, $B$, $C$, and $D$) are running in a qualifying race (assume that there are no ties). The top two finishers qualify for the finals. The observation is the pair of runners that qualify for the finals.

5. The board of directors of Fibber Corporation has five members ($A$, $B$, $C$, $D$, and $E$). Using set notation write out the sample space for each of the following random experiments:
   (a) A chairman and a treasurer are elected.
   (b) Three directors are selected to form a search committee to hire a new CEO.

6. Using set notation, write out the sample space for each of the following random experiments:
   (a) You reach into a large jar containing jelly beans of four different flavors [Juicy Pear ($J$), Kiwi ($K$), Licorice ($L$), and Mango ($M$)] and grab two jelly beans at random. The observation is the flavors of the two jelly beans.
   (b) One of the playlists on your iPod contains songs (more than one) from each of your five favorite artists: Adele ($A$), Bruce Springsteen ($B$), Carrie Underwood ($C$), Dave Matthews ($D$), and Elvis Costello ($E$). Suppose that two songs are randomly chosen from the playlist and played. The observation is the set of artists you get to hear (regardless of order).

*In Exercises 7 through 10, the sample spaces are too big to write down in full. In these exercises, you should describe the sample space either by describing a generic outcome or by listing some outcomes and then using the . . . notation. In the latter case, you should write down enough outcomes to make the description reasonably clear.*

7. A coin is tossed 10 times in a row. The observation is how the coin lands ($H$ or $T$) on each toss. Describe the sample space.

8. A student randomly guesses the answers to a 10-question true-or-false quiz. The observation is the student's answer ($T$ or $F$) for each question. Describe the sample space.

9. A die is rolled four times in a row. The observation is the number that comes up on each roll. Describe the sample space.

10. A student randomly guesses the answers to a multiple-choice quiz consisting of 10 questions. The observation is the student's answer ($A$, $B$, $C$, $D$, or $E$) for each question. Describe the sample space.

11. A coin is tossed three times in a row. The observation is how the coin lands (heads or tails) on each toss [see Exercise 1(a)]. Write out the event described by each of the following statements as a set.
   (a) $E_1$: "the coin comes up heads exactly twice."
   (b) $E_2$: "all three tosses come up the same."
   (c) $E_3$: "exactly half of the tosses come up heads."
   (d) $E_4$: "the first two tosses come up tails."

12. A student randomly guesses the answers to a four-question true-or-false quiz. The observation is the student's answer (*T* or *F*) for each question [see Exercise 2(b)]. Write out the event described by each of the following statements as a set.

(a) $E_1$: "the student answers $T$ to two out of the four questions."

(b) $E_2$: "the student answers $T$ to *at least* two out of the four questions."

(c) $E_3$: "the student answers $T$ to *at most* two out of the four questions."

(d) $E_4$: "the student answers $T$ to the first two questions."

13. A pair of dice is rolled. The observation is the number that comes up on each die (see Example 16.4). Write out the event described by each of the following statements as a set.

(a) $E_1$: "*pairs* are rolled." (A "pair" is a roll in which both dice come up the same number.)

(b) $E_2$: "*craps* are rolled." ("Craps" is a roll in which the sum of the two numbers rolled is 2, 3, or 12.)

(c) $E_3$: "a *natural* is rolled." (A "natural" is a roll in which the sum of the two numbers rolled is 7 or 11.)

14. A card is drawn out of a standard deck of 52 cards. Each card can be described by giving its "value" ($A, 2, 3, 4, \ldots, 10, J, Q, K$) and its "suit" ($H$ for hearts, $C$ for clubs, $D$ for diamonds, and $S$ for spades). For example, $2D$ denotes the two of diamonds and $JH$ denotes the jack of hearts. Write out the event described by each of the following statements as a set.

(a) $E_1$: "draw a queen."

(b) $E_2$: "draw a heart."

(c) $E_3$: "draw a *face* card." (A "face" card is a jack, queen, or king.)

15. A coin is tossed 10 times in a row. The observation is how the coin lands (*H* or *T*) on each toss (see Exercise 7). Write out the event described by each of the following statements as a set.

(a) $E_1$: "none of the tosses comes up tails."

(b) $E_2$: "exactly one of the 10 tosses comes up tails."

(c) $E_3$: "nine or more of the tosses come up tails."

16. Five candidates ($A$, $B$, $C$, $D$, and $E$) have a chance to be selected to be on *American Idol*. Any subset of them (including none of them or all of them) can be selected. The observation is which subset of individuals is selected. Write out the event described by each of the following statements as a set.

(a) $E_1$: "two candidates get selected."

(b) $E_2$: "three candidates get selected."

(c) $E_3$: "three candidates get selected, and $A$ is not one of them."

## 16.2 The Multiplication Rule, Permutations, and Combinations

17. A California license plate starts with a digit other than 0 followed by three capital letters followed by three more digits (0 through 9).

(a) How many different California license plates are possible?

(b) How many different California license plates start with a 5 and end with a 9?

(c) How many different California license plates have no repeated symbols (all the digits are different and all the letters are different)?

18. A computer password consists of four letters ($A$ through $Z$) followed by a single digit (0 through 9). Assume that the passwords are not case sensitive (i.e., that an uppercase letter is the same as a lowercase letter).

(a) How many different passwords are possible?

(b) How many different passwords end in 1?

(c) How many different passwords do not start with $Z$?

(d) How many different passwords have no $Z$'s in them?

19. Jack packs two pairs of shoes, one pair of boots, three pairs of jeans, four pairs of dress pants, and three dress shirts for a trip.

(a) How many different outfits can Jack make with these items?

(b) If Jack were to bring along two sweaters so that he could wear either a dress shirt or a sweater or he could wear both a dress shirt and a sweater, how many outfits could Jack make?

20. A French restaurant offers a menu consisting of three different appetizers, two different soups, four different salads, nine different main courses, and five different desserts.

(a) A fixed-price lunch meal consists of a choice of appetizer, salad, and main course. How many different fixed-price lunch meals are possible?

(b) A fixed-price dinner meal consists of a choice of appetizer, a choice of soup or salad, a main course, and a dessert. How many different fixed-price dinner meals are possible?

(c) A dinner special consists of a choice of soup, salad, or both, plus a main course. How many dinner specials are possible?

21. A set of reference books consists of eight volumes numbered 1 through 8.

(a) In how many ways can the eight books be arranged on a shelf?

(b) In how many ways can the eight books be arranged on a shelf so that at least one book is out of order?

**22.** Nine people (four men and five women) line up at a checkout stand in a grocery store.

(a) In how many ways can they line up?

(b) In how many ways can they line up if the first person in line must be a woman?

(c) In how many ways can they line up if they must alternate woman, man, woman, man, and so on?

**23.** How many seven-digit numbers (i.e., numbers from 1,000,000 through 9,999,999)

(a) are even?

(b) are divisible by 5?

(c) are divisible by 25?

**24.** How many nine-digit numbers (i.e., numbers between 100,000,000 and 999,999,999)

(a) have no repeated digits?

(b) are number palindromes? (A *number palindrome* is a number such as 374323473 that reads the same whether read from left to right or from right to left.)

**25.** Determine the number of outcomes $N$ in each sample space.

(a) A coin is tossed 10 times in a row. The result of each toss ($H$ or $T$) is observed.

(b) A die is rolled four times in a row. The number that comes up on each roll is observed.

(c) A die is rolled four times in a row. The sum of the numbers rolled is observed.

**26.** Determine the number of outcomes $N$ in each sample space.

(a) A student randomly answers a 10-question true-false quiz. The student's answer ($T$ or $F$) to each question is observed.

(b) A student randomly answers a 10-question multiple-choice quiz. The student's answer ($A$, $B$, $C$, $D$, or $E$) to each question is observed.

*In Exercises 27 through 30, you are asked to give your answer using the notation ${}_nP_r$ or ${}_nC_r$. You do not have to give an actual numerical answer, so you won't need a calculator.*

**27.** The board of directors of the XYZ Corporation has 15 members.

(a) How many different slates of four officers (a President, a Vice President, a Treasurer, and a Secretary) can be chosen?

(b) A four-person committee needs to be selected to conduct a search for a new CEO. In how many ways can the search committee be selected?

**28.** There are 10 athletes entered in an Olympic event.

(a) In how many ways can one pick the winners of the gold, silver, and bronze medals?

(b) In how many ways can one pick the seven athletes who will not earn any medals?

**29.** The Brute Force Bandits is a punk rock band planning their next concert tour. The band has a total of 30 new songs in their repertoire.

(a) How many different set lists of 18 songs can the band select to play on the tour? (*Hint*: Assume that the order in which the songs are listed is irrelevant.)

(b) How many different ways are there for the band to record a CD consisting of 18 songs chosen from the 30 new songs? (*Hint*: In recording a CD, the order in which the songs appear on the CD is relevant.)

**30.** There are 347 NCAA Division I college basketball teams.

(a) How many different top-25 rankings are possible? [Assume that every team has a chance to be a top-25 team.]

(b) How many ways are there to choose 64 teams (unseeded) to make it to the NCAA tournament? [Assume every combination of 64 teams is possible.]

**31.** A major league baseball team roster consists of 40 players of which 25 are considered active.

(a) How many ways are there for a manager to select 25 active players from a major league roster?

(b) How many ways can a manager select a nine-player batting lineup from the active roster for opening day?

**32.** Bob has 20 different dress shirts in his wardrobe.

(a) In how many ways can Bob select seven shirts to pack for a business trip?

(b) In how many ways can Bob select 5 of the 7 dress shirts he packed for the business trip—one for the Monday meeting, one for the Tuesday dinner, one for the Wednesday party, one for the Thursday conference, and one for the Friday date?

## 16.3 Probabilities and Odds

**33.** Consider the sample space $S = \{o_1, o_2, o_3, o_4, o_5\}$. Suppose that $\Pr(o_1) = 0.22$ and $\Pr(o_2) = 0.24$.

(a) Find the probability assignment for the probability space when $o_3$, $o_4$, and $o_5$ all have the same probability.

(b) Find the probability assignment for the probability space when $\Pr(o_5) = 0.1$ and $o_3$ has the same probability as $o_4$ and $o_5$ combined.

**34.** Consider the sample space $S = \{o_1, o_2, o_3, o_4\}$. Suppose that $\Pr(o_1) + \Pr(o_2) = \Pr(o_3) + \Pr(o_4)$ and that $\Pr(o_1) = 0.15$.

(a) Find the probability assignment for the probability space when $o_2$ and $o_3$ have the same probability.

(b) Find the probability assignment for the probability space when $\Pr(o_3) = 0.22$.

**35.** Four candidates are running for mayor of Happyville. According to the polls candidate $A$ has a "one in five" probability of winning [i.e., $\Pr(A) = 1/5$]. Of the other three candidates, all we know is that candidate $C$ is twice as likely to win as candidate $B$ and that candidate $D$ is three times as likely to win as candidate $B$. Find the probability assignment for this probability space.

36. Seven horses ($A$, $B$, $C$, $D$, $E$, $F$, and $G$) are running in the Boonsville Sweepstakes. According to the oddsmakers, $A$ has a "one in four" probability of winning [i.e., $\Pr(A) = 1/4$], $B$ has a "three in ten" probability of winning, and $C$ has a "one in twenty" probability of winning. The remaining four horses all have the same probability of winning. Find the probability assignment for the probability space.

37. The circular spinner shown in Fig. 16-7 has a red sector with a central angle of 108°, a blue sector and a white sector each with central angles of 72°, and a green sector and a yellow sector having the same central angle. Assume that the needle is spun so that it randomly stops at one of these sectors (if it stops exactly on the line between sectors, then the needle is spun again). Find the probability assignment for the probability space.

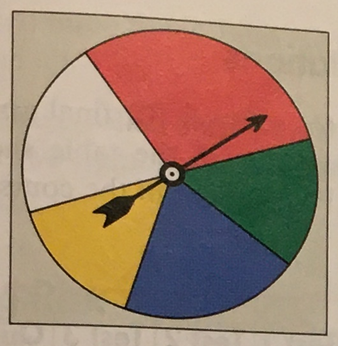


FIGURE 16-7

38. The circular spinner shown in Figure 16-8 is divided into four sectors. Sectors A and B have the same size. Sectors C and D are each twice as large as sector A. Assume that the needle is spun so that it randomly stops at one of these sectors (if it stops on the line between sectors, then the needle is spun again). Find the probability assignment for the probability space.

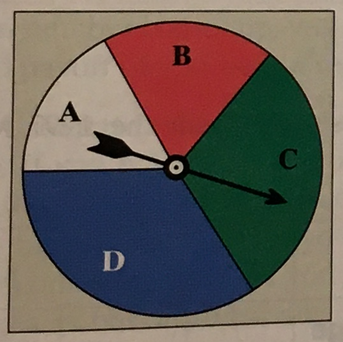



FIGURE 16-8

39. An honest coin is tossed three times in a row. Find the probability of each of the following events. (*Hint*: Do Exercise 11 first.)

(a) $E_1$: "the coin comes up heads exactly twice."

(b) $E_2$: "all three tosses come up the same."

(c) $E_3$: "exactly half of the tosses come up heads."

(d) $E_4$: "the first two tosses come up tails."

40. A student randomly guesses the answers to a four-question true-or-false ($T - F$) quiz. Find the probability of each of the following events. (*Hint:* Do Exercise 12 first.)

(a) $E_1$: "the student answers $F$ on two of the four questions."

(b) $E_2$: "the student answers $F$ on *at least* two of the four questions."

(c) $E_3$: "the student answers $F$ on *at most* two of the four questions."

(d) $E_4$: "the student answers $F$ to the first two questions."

41. A pair of honest dice is rolled. Find the probability of each of the following events. (*Hint:* Do Exercise 13 first.)

(a) $E_1$: "*pairs* are rolled." (A "pair" is a roll in which both dice come up the same number.)

(b) $E_2$: "*craps* are rolled." ("Craps" is a roll in which the sum of the two numbers rolled is 2, 3, or 12.)

(c) $E_3$: "a *natural* is rolled." (A "natural" is a roll in which the sum of the two numbers rolled is 7 or 11.)

42. A card is drawn at random out of a well-shuffled deck of 52 cards. Find the probability of each of the following events. (*Hint*: Do Exercise 14 first.)

(a) $E_1$: "draw a queen."

(b) $E_2$: "draw a heart."

(c) $E_3$: "draw a *face* card." (A "face" card is a jack, queen, or king.)

43. An honest coin is tossed 10 times in a row. Find the probability of each of the following events. (*Hint*: Do Exercise 15 first.)

(a) $E_1$: "none of the tosses comes up tails."

(b) $E_2$: "exactly one of the 10 tosses comes up tails."

(c) $E_3$: "nine or more of the tosses come up tails."

44. Five candidates ($A$, $B$, $C$, $D$, and $E$) have a chance to be selected to be on *American Idol*. Any subset of them (including none of them or all of them) can be selected, and assume that the selection process is completely random (the subsets of candidates are all equally likely). Find the probability of each of the following events. (*Hint*: Do Exercise 16 first.)

(a) $E_1$: "two candidates get selected."

(b) $E_2$: "three candidates get selected."

(c) $E_3$: "three candidates get selected, and $A$ is not one of them."

45. A student takes a 10-question true-or-false quiz and randomly guesses the answer to each question. Suppose that a correct answer is worth 1 point and an incorrect answer is worth −0.5 points. Find the probability that the student

(a) gets 10 points.

(b) gets −5 points.

(c) gets 8.5 points.

(d) gets 8 or more points.

(e) gets 5 points.

(f) gets 7 or more points.

**46.** Suppose that the probability of giving birth to a boy and the probability of giving birth to a girl are both 0.5. Find the probability that in a family of four children,

(a) all four children are girls.

(b) there are two girls and two boys.

(c) the youngest child is a girl.

(d) the oldest child is a boy.

**47.** The Tasmania State University glee club has 15 members. A *quartet* of four members must be chosen to sing at the university president's reception. Assume that the quartet is chosen randomly by drawing the names out of a hat. Find the probability that

(a) Alice (one of the members of the glee club) is chosen to be in the quartet.

(b) Alice is not chosen to be in the quartet.

(c) the four members chosen for the quartet are Alice, Bert, Cathy, and Dale.

**48.** Ten professional basketball teams are participating in a draft lottery. (A draft lottery is a lottery to determine the order in which teams get to draft players.) Ten balls, each containing the name of one team (call them $A$, $B$, $C$, $D$, $E$, $F$, $G$, $H$, $I$, and $J$ for short), are placed in an urn and thoroughly mixed. Four balls are drawn, one at a time, from the urn. The four teams chosen get to draft first and in the order they are chosen. The remaining six teams have to draft in reverse order of season records. Find the probability that

(a) $A$ is the first team chosen.

(b) $A$ is one of the four teams chosen.

(c) $A$ is not one of the four teams chosen.

**49.** An honest coin is tossed 10 times in a row. The result of each toss ($H$ or $T$) is observed. Find the probability of the event $E =$ "a $T$ comes up at least once." (*Hint:* Find the probability of the complementary event.)

**50.** Imagine a game in which you roll an honest die three times. Find the probability of the event $E =$ "at least one of the rolls of the dice comes up a 6." (*Hint:* See Example 16.22.)

**51.** Find the odds of each of the following events.

(a) an event $E$ with $\Pr(E) = 4/7$

(b) an event $E$ with $\Pr(E) = 0.6$

**52.** Find the odds of each of the following events.

(a) an event $E$ with $\Pr(E) = 3/11$

(b) an event $E$ with $\Pr(E) = 0.375$

**53.** In each case, find the probability of an event $E$ having the given odds.

(a) The odds in favor of $E$ are 3 to 5.

(b) The odds against $E$ are 8 to 15.

**54.** In each case, find the probability of an event $E$ having the given odds.

(a) The odds in favor of $E$ are 4 to 3.

(b) The odds against $E$ are 12 to 5.

(c) The odds in favor of $E$ are the same as the odds against $E$.

**55.** There are three possible things that can happen over the next year to real estate values in the Sunnyside neighborhood—they can go up, they can go down, or they can remain the same. The odds that real estate values will go up are 1 to 1, and the odds that they will go down are 2 to 3. What are the odds that real estate values will remain the same?

**56.** Three candidates—Aguilera, Bieber, and Cyrus—are running for mayor of Kingsburg. The odds of Aguilera winning are 1 to 2, and the odds of Cyrus winning are 2 to 7. What are the odds of Bieber winning?

## 16.4 Expectations

**57.** The scoring for a Psych 101 final grade is shown in Table 16-18. The last row of the table shows Paul's individual scores. Find Paul's score in the course, expressed as a percent.

| | Test 1 | Test 2 | Test 3 | Quizzes | Paper | Final |
|---|---|---|---|---|---|---|
| Weight | 15% | 15% | 15% | 10% | 25% | 20% |
| Points possible | 100 | 100 | 100 | 120 | 100 | 180 |
| Paul's score | 77 | 83 | 91 | 90 | 87 | 144 |

■ TABLE 16-18

**58.** Tiger Woods's scores in the 1997 Masters are shown in Table 16-19.

| Score | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Percentage of holes | 1.4% | 36.1% | 50% | 11.1% | 1.4% |

■ TABLE 16-19

What was Tiger Woods's average score per hole?

**59.** At Thomas Jefferson High School, the student body is divided by age as follows: 7% of the students are 14, 22% of the students are 15, 24% of the students are 16, 23% of the students are 17, 19% of the students are 18, and the rest of the students are 19. Find the average age of the students at Thomas Jefferson High School.

**60.** In 2005 the Middletown Zoo averaged 4000 visitors on sunny days, 3000 visitors on cloudy days, 1500 visitors on rainy days, and only 100 visitors on snowy days. The percentage of days of each type in 2005 shown in Table 16-20. Find the average daily attendance at the Middletown Zoo for 2005.

# 16.1 Sample Spaces and Events

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

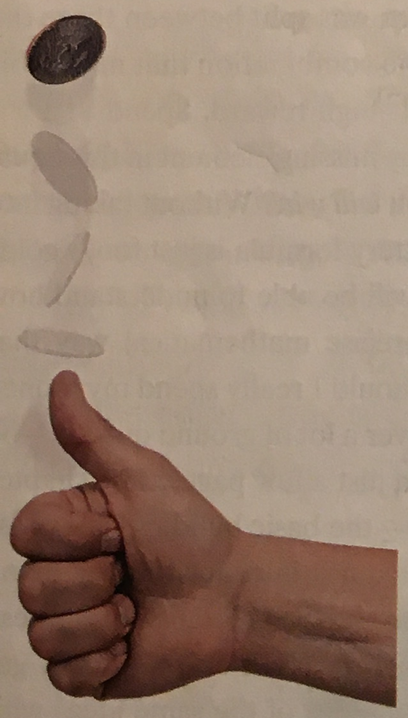
We will use the term **random experiment** to describe an activity or a process whose outcome cannot be predicted ahead of time. Typical examples of random experiments are tossing a coin, rolling a pair of dice, drawing cards out of a deck of cards, predicting the result of a football game, and forecasting 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 random experiments for which there is only a finite set of outcomes, although sample spaces with infinitely many outcomes are both possible and important.

We illustrate the concept of a sample space by means of several examples. Since the sample space of any experiment is a set of individual 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 16.1 TOSSING A COIN

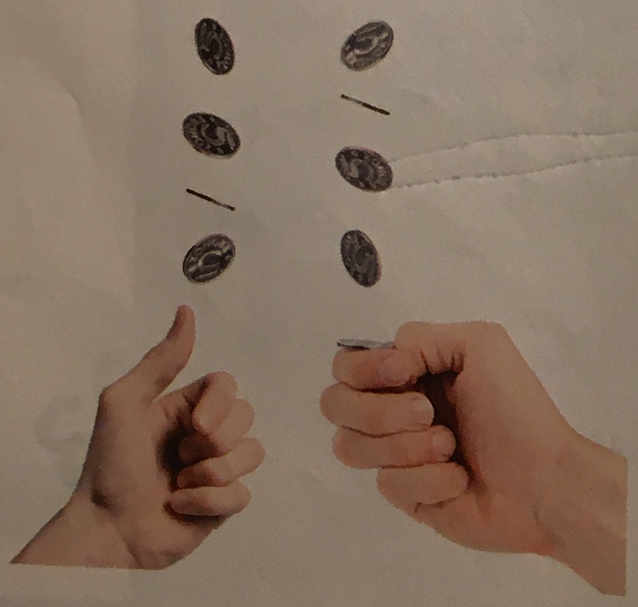
One simple random experiment is to *toss a quarter* and *observe whether it lands heads or tails*. The sample space can be described by $S = \{H, T\}$, where $H$ stands for *Heads* and $T$ for *Tails*. Here $N = 2$.



A couple of comments about coins are in order here. First, the fact that the coin in Example 16.1 is 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 where 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 16.2 TOSSING A COIN TWICE: EPISODE 1

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 coins* 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 coin tossed by the left hand and the second the coin tossed by the right hand.)

Since they have the same sample space, we will consider the two random experiments just described as equivalent, so whether you toss two coins at once or one coin twice, from our point of view there is no difference between the two.

| Weather condition | Sunny | Cloudy | Rainy | Snowy |
|---|---|---|---|---|
| Percentage of days | 47% | 27% | 19% | 7% |

■ TABLE 16-20

61. A box contains twenty \$1 bills, ten \$5 bills, five \$10 bills, four \$20 bills, and one \$100 bill. You blindly reach into the box and draw a bill at random. What is the expected value of your draw?

62. A basketball player shoots two consecutive free throws. Each free-throw is worth 1 point and has probability of success $p = 3/4$. Let $X$ denote the number of points scored. Find the expected value of $X$.

63. A fair coin is tossed three times. Find the expected number of heads that come up.

64. A pair of honest dice is rolled once. Find the expected value of the sum of the two numbers rolled. (*Hint*: See Example 16.20).

## 16.5 Measuring Risk

65. Suppose that you roll a pair of honest dice. If you roll a total of 7, you win \$18; if you roll a total of 11, you win \$54; if you roll any other total, you lose \$9. Find the expected payoff for this game.

66. On an American roulette wheel, there are 18 red numbers, 18 black numbers, plus 2 green numbers (0 and 00). If you bet \$$N$ on red, you win \$$N$ if a red number comes up (i.e., you get \$$2N$ back—your original bet plus your \$$N$ profit); if a black or green number comes up, you lose your \$$N$ bet. Find the expected payoff of a \$1 bet on red.

67. On an American roulette wheel, there are 38 numbers: 00, 0, 1, 2, . . . , 36. If you bet \$$N$ on any one number—say, for example, on 10—you win \$$36N$ if 10 comes up (i.e., you get \$$37N$ back—your original bet plus your \$$36N$ profit); if any other number comes up, you lose your \$$N$ bet. Find the expected payoff of a \$1 bet on 10 (or any other number).

68. Suppose that you roll a single die. If an odd number (1, 3, or 5) comes up, you win the amount of your roll (\$1, \$3, or \$5, respectively). If an even number (2, 4, or 6) comes up, you have to pay the house the amount of your roll (\$2, \$4, or \$6, respectively).

    (a) Find the expected payoff for this game.

    (b) Is this a fair game? Explain.

69. Joe is buying a new plasma TV at Circuit Town. The salesman offers Joe a three-year extended warranty for \$80. The salesman tells Joe that 24% of these plasma TVs require repairs within the first three years, and the average cost of a repair is \$400. Should Joe buy the extended warranty? Explain your reasoning.

70. Jackie is buying a new MP3 player from Better Buy. The store offers her a two-year extended warranty for \$19. Jackie read in a consumer magazine that for this model MP3, 5% require repairs within the first two years at an average cost of \$50. Should Jackie buy the extended warranty? Explain your reasoning.

71. The service history of the Prego SUV is as follows: 50% will need no repairs during the first year, 35% will have repair costs of \$500 during the first year, 12% will have repair costs of \$1500 during the first year, and the remaining SUVs (the real lemons) will have repair costs of \$4000 during their first year. Determine the price that the insurance company should charge for a one-year extended warranty on a Prego SUV if it wants to make an average profit of \$50 per policy.

72. An insurance company plans to sell a \$250,000 one-year term life insurance policy to a 60-year-old male. Of 2.5 million men having similar risk factors, the company estimates that 7500 of them will die in the next year. What is the premium that the insurance company should charge if it would like to make a profit of \$50 on each policy?

## JOGGING

73. The ski club at Tasmania State University has 35 members (15 females and 20 males). A committee of three members—a President, a Vice President, and a Treasurer—must be chosen.

    (a) How many different three-member committees can be chosen?

    (b) How many different three-member committees can be chosen in which the committee members are all females?

    (c) How many different three-member committees can be chosen in which the committee members are all the same gender?

    (d) How many different three-member committees can be chosen in which the committee members are *not* all the same gender?

74. The ski club at Tasmania State University has 35 members (15 females and 20 males). A committee of four members of equal standing must be chosen.

    (a) How many different four-member committees can be chosen?

    (b) How many different four-member committees can be chosen consisting of two males and two females?

75. Andy and Roger are playing in a tennis match. (A tennis match is a best-of-five contest: The first player to win three games wins the match, and there are no ties.) We can describe the outcome of the tennis match by a string of letters ($A$ or $R$) that indicate the winner of each game. For example, the string $RARR$ represents an outcome where Roger wins games 1, 3, and 4, at which point the match is over (game 5 is not played).

    (a) Describe the event "Roger wins the match in game 5."

    (b) Describe the event "Roger wins the match."

    (c) Describe the event "the match goes five games."

**76.** Two teams (call them $X$ and $Y$) play against each other in the World Series. (The World Series is a best-of-seven series: The first team to win four games wins the series, and games cannot end in a tie.) We can describe an outcome for the World Series by writing a string of letters that indicate (in order) the winner of each game. For example, the string $XYXXYX$ represents the outcome: $X$ wins game 1, $Y$ wins game 2, $X$ wins game 3, and so on.

**(a)** Describe the event "$X$ wins in five games."

**(b)** Describe the event "the series is over in game 5."

**(c)** Describe the event "the series is over in game 6."

**(d)** Find the size of the sample space.

**77.** A study group of 15 students is to be split into three groups of five students each. In how many ways can this be done?

**78.** Given that ${}_{150}P_{50} \approx 6.12 \times 10^{104}$, find an approximate value for ${}_{150}P_{51}$.

**79.** Given that ${}_{150}C_{50} \approx 2.01 \times 10^{40}$, find an approximate value for ${}_{150}C_{51}$.

**80.** An urn contains seven red balls and three blue balls.

**(a)** If three balls are selected all at once, what is the probability that two are blue and one is red?

**(b)** If three balls are selected by pulling out a ball, noting its color, and putting it back in the urn before the next selection, what is the probability that only the first and third balls drawn are blue?

**(c)** If three balls are selected one at a time without putting them back in the urn, what is the probability that only the first and third balls drawn are blue?

**81.** If we toss an honest coin 10 times, what is the probability of

**(a)** getting 5 heads and 5 tails?

**(b)** getting 3 heads and 7 tails?

**82.** A draw poker hand consists of 5 cards taken from a deck of 52 cards where the order of the cards is irrelevant (see Examples 16.15 and 16.23). Assuming you are playing with an honest deck, find the probability of each of the following hands.

**(a)** A "four of a kind" (four cards of the same value and a "kicker" of any other value)

**(b)** A "flush" (all five cards of the same suit)

**(c)** An "ace-high straight" (10, J, Q, K, A of any suit but not all the same suit)

**(d)** A "full house" (three cards of equal value and two other cards of equal value)

**83.** There are 500 tickets sold in a raffle. If you have three of these tickets and five prizes are to be given, what is the probability that you will win at least one prize? (Give your answer in symbolic notation.)

**84.** This exercise refers to the game of chuck-a-luck discussed in Example 16.34. Explain why, when you roll three dice,

**(a)** the probability of rolling two 4's plus another number (not a 4) is 15/216.

**(b)** the probability of rolling one 4 plus two other numbers (not 4's) is 75/216.

**(c)** the probability of rolling no 4's is 125/216.

**85.** In head-to-head, 7-card stud poker you make your hand by selecting your 5 best cards from the 2 in your hand and 3 from the 5 common cards showing on the table (the "board"). Suppose you are holding A♠, 2♠, in your hand and Q♠, K♥, 3♠, 7♦, are showing on the board (this is called a "flush draw")—if the last card ("river" card) is a spade you will have an ace high flush and a guaranteed win. Assume that your opponent has a decent hand and if you don't get the spade on the river card you will lose the hand.

**(a)** Suppose there is \$100 in the pot and your opponent moves "all-in" with a \$50 bet. Should you call the bet or fold? Explain.

**(b)** Suppose there is \$100 in the pot and your opponent moves "all-in" with a \$20 bet. Should you call the bet or fold? Explain.

## RUNNING

**86.** If an honest coin is tossed $N$ times, what is the probability of getting the same number of heads as tails? (*Hint*: Consider two cases: $N$ even and $N$ odd.)

**87.** How many different "words" (they don't have to mean anything) can be formed using all the letters in

**(a)** the word PARSLEY.

**(b)** the word PEPPER. [*Hint*: This question is very different (and harder) than the one in (a). What's the difference?]

**88. Shooting craps.** In the game of craps, the player's first roll of a pair of dice is very important. If the first roll is 7 or 11, the player wins. If the first roll is 2, 3, or 12, the player loses. If the first roll is any other number (4, 5, 6, 8, 9, 10), this number is called the player's "point." The player then continues to roll until either the point reappears, in which case the player wins, or until a 7 shows up before the point, in which case the player loses. What is the probability that the player will win? (Assume that the dice are honest.)

**89. The birthday problem.** There are 30 people in a room. What is the probability that at least two of these people have the same birthday—that is, have their birthdays on the same day and month?

90. **Yahtzee.** Yahtzee is a dice game in which five standard dice are rolled at one time.

(a) What is the probability of scoring "Yahtzee" with one roll of the dice? (You score Yahtzee when all five dice match.)

(b) What is the probability of a *four of a kind* with one roll of the dice? (A *four of a kind* is rolled when four of the five dice match.)

(c) What is the probability of rolling a *large straight* in one roll of the dice? (A *large straight* consists of five numbers in succession on the dice.)

(d) What is the probability of rolling *trips* with one roll of the dice? (*Trips* are rolled when three of the five dice match and the other two dice do not match the trips or each other.)

## PROJECTS AND PAPERS

### 1 A History of Gambling

Games of chance, be they for religious purposes or for profit, can be traced all the way back to some of the earliest civilizations—the Babylonians, Assyrians, and ancient Egyptians were all known to play primitive dice games as well as board games of various kinds.

In this project you are to write a paper discussing the history of gambling and the historical role of games of chance in the development of probability theory.

### 2 The Letters Between Pascal and Fermat

In the summer of 1654, Pierre de Fermat and Blaise Pascal—two of the most famous mathematicians of their time—exchanged a series of famous letters discussing certain mathematical problems related to gambling. The problems were originally proposed to Pascal by his friend, Chevalier de Méré—a nobleman and notorious gambler with a flair for raising interesting mathematical questions. The exchange of letters between Pascal and Fermat is of particular interest because it laid the foundation for the future development of the mathematical theory of probability. While not all of the correspondence between Fermat and Pascal has survived, enough of the letters exist to see how both mathematicians approached, each in his own way, de Méré's questions.

In this project, you are to write a paper discussing the famous exchange of letters between Pascal and Fermat. Describe the questions that were discussed and the approach that each mathematician took to solve them.

### 3 Book Review: *Conned Again, Watson! Cautionary Tales of Logic, Math, and Probability*

As its title suggests, *Conned Again, Watson! Cautionary Tales of Logic, Math, and Probability* by Colin Bruce (Perseus Publishing, 2002) is a collection of mathematical tales set in the form of Sherlock Holmes adventures. Each of the stories has a mathematical connection, and several touch on issues related to the theme of this chapter (particularly relevant are "The Case of the Gambling Nobleman" and "The Case of the Surprise Heir," but several others deal with probability topics as well).

Choose one of the stories in the book with a probabilistic connection and write a review of the story. Pay special attention to the mathematical issues raised in the story.

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 16.2 is that a random experiment is defined by two things: what the act is (such as tossing coins) and what the observation that follows the act is.

**EXAMPLE 16.3** **SHOOTING FREE THROWS: EPISODE 1**

Here is a familiar scenario: Your favorite basketball team is down by 1, clock running out, and one of the players in your team 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 16.2, the answer depends on a few subtleties.

In one scenario (the *penalty situation*) the 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 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 the player is shooting a *one-and-one*. This means that the player gets to shoot the second free throw only if he makes the first one. In this case there are also three possible outcomes, but the circumstances are different because the sequence 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 16.4** **ROLLING A PAIR OF DICE: EPISODE 1**

The most general scenario when rolling a pair of dice is to consider each die separately (in this regard, it helps to think of them as being of different colors—say one is white and the other one is red). The sample space consists of the 36 outcomes shown in Fig. 16-1.

Throughout this chapter we will see several examples of rolling a pair of dice, so for convenience we will use the notation $(1, 1), (1, 2), \ldots, (6, 6)$ to denote the outcome of a roll. The first number represents one die (say white) and the second number the other die (say red).

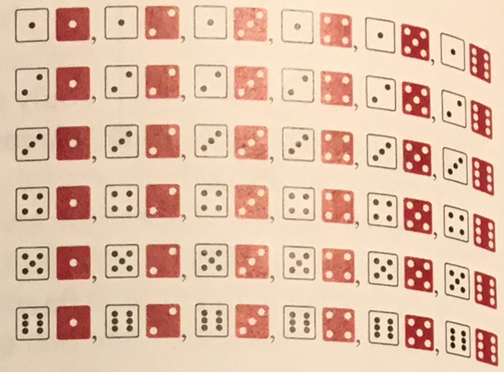


FIGURE 16-1

An alternative 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—$(1,6)$, $(2,5)$, $(3,4)$, $(4,3)$, $(5,2)$ and $(6,1)$—but 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 16.5 A CAR AND TWO GOATS: EPISODE 1

Imagine you are a contestant in a television game show. The host of the show shows you three doors (we'll call them Door #1, Door #2 and Door #3). You get to pick one of the three doors and keep the prize hidden behind it—a fancy new sports car behind one of the three doors and a goat behind each of the other two. (If this story sounds familiar to you, then you are not too young to remember Monty Hall and the long-running game show *Let's Make a Deal*.)

This game gives us a slightly more exotic example of a random experiment than the more mundane tossing of a pair of dice. The randomness comes from the fact that the sports car is randomly placed behind one of the three doors and you must guess which one (we are assuming here that you want to win the sports car and you are not too interested in taking home a goat).

As you blindly confront the decision on which door to choose you think of sample spaces. What is the proper sample space that describes your predicament? If we use the notation $(x, y)$ to denote the fact that you choose Door #$x$ and the car is behind Door #$y$, we can describe the sample space as consisting of the following nine outcomes:

$$S = \{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3)\}.$$

Of the above, $(1,1)$, $(2,2)$ and $(3,3)$ mean you are going home driving a fancy new sports car, the other six mean you get a goat. (In the *Let's Make a Deal* show the undesirable prizes were called Zonks, and participants could trade the Zonk for a more practical consolation prize like a toaster or a television set.)

Imagine now that you make your decision and choose one of the three doors—say you choose Door #1. The host, who knows where the car is, opens Door #3 to show there is a goat behind it and offers you the option of switching. Should you stay with your original choice or switch to Door #2? This is an interesting question, and we will return to it later in the chapter.

**EXAMPLE 16.6** **HORSE RACES AND TRIFECTAS: EPISODE 1**

In the language of horse racing, a *trifecta* is a bet on predicting the first-, second-, and third-place finishers in a horse race in the right order. A trifecta ticket has boxes for the first-, second-, and third-place finishers in which you can enter the numbers of the three horses you predict will finish first, second, and third, respectively.

Imagine you bought a trifecta ticket in a horse race with 10 horses numbered 1 through 10. You marked horse number 2 to finish first, number 6 to finish second, and number 5 to finish third. We will use the notation (2, 6, 5) to denote your trifecta ticket. Note that if horse number 6 finishes first, horse number 2 second, and horse number 5 third you are out of luck—close but no cigar.

The random experiment in this example is the running of the race, and the observation of interest is which horses finished first, second, and third, respectively. The sample space consists of all possible outcomes of the form $(x, y, z)$ where $x$, $y$, and $z$, are three different numbers between 1 and 10. There are 720 outcomes in this sample space (we will see why this is the case in Example 16.12), so writing them all down, one by one, is not such a good idea. There are a couple of options, the simplest one is using the "..." notation, as in $S = \{(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), \ldots, (10,9, 8)\}$. The "..." is a way of saying "and so on," somewhat along the same lines as the "..." notation used in Chapter 9 for describing sequences. ■

Example 16.6 illustrates the point that sample spaces can be quite large and have a lot of outcomes and that we are often justified in not writing each and every one of these outcomes down. This is where the "..." comes in handy. The key thing is to understand what the sample space looks like and most importantly, 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 section.

## Events

An **event** is a *subset* of the sample space, and any *subset* of the sample space can be considered an event. By definition, events are sets, so we will use set notation to describe them and the basic rules of set theory to work with them.

In Section 2.4 we introduced the fact that a subset with $N$ elements has $2^N$ subsets—it follows that in a sample space with $N$ outcomes there are $2^N$ different events.

**EXAMPLE 16.7** **TOSSING A COIN TWICE: EPISODE TWO**

If you toss a coin twice and record the result of each toss, the sample space is $S = \{HH, HT, TH, TT\}$—we discussed this in Example 16.2. This sample space has 16 different subsets, each representing a different event, as shown in Table 16-1.

| | | | |
|---|---|---|---|
| 1. $\{\ \}$ | 5. $\{TT\}$ | 9. $\{HT, TH\}$ | 13. $\{HH, HT, TT\}$ |
| 2. $\{HH\}$ | 6. $\{HH, HT\}$ | 10. $\{HT, TT\}$ | 14. $\{HH, TH, TT\}$ |
| 3. $\{HT\}$ | 7. $\{HH, TH\}$ | 11. $\{TH, TT\}$ | 15. $\{HT, TH, TT\}$ |
| 4. $\{TH\}$ | 8. $\{HH, TT\}$ | 12. $\{HH, HT, TH\}$ | 16. $\{HH, HT, TH, TT\}$ |

■ **TABLE 16-1** The 16 possible events when tossing a coin twice.

Several of the events listed in Table 16-1 deserve special attention.

- The empty set $\{\ \}$ is considered an event, essentially an event with no outcomes. At first glance, the idea of an event with no outcomes might seem a little weird, but we will soon see that it is very convenient to have it in our toolkit. The event with no outcomes is called the **impossible event**.
- The entire sample space $S$ is an event, the event where every possible outcome is included. When viewed as an event, the entire sample space is called the **certain event**.
- The events $\{HH\}$, $\{HT\}$, $\{TH\}$, and $\{TT\}$ consist of just one outcome. An event consisting of a single outcome is called a **simple event**. Simple events are just the individual outcomes in a sample space reinterpreted in the language of events.

A convenient way to think of an event is as a "package" in which outcomes with some common characteristic are bundled together, but the "package" idea has to be interpreted very broadly: The package can have nothing in it, everything in it, or just a single item in it.

### EXAMPLE 16.8 TOSSING A COIN THREE TIMES

If you toss a coin three consecutive times and record the result of each toss, the sample space is $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$, $N = 8$, and the number of possible events is $2^8 = 256$—too many to list. Table 16-2 shows a few of these events, described both in plain English and in the language of sets.

| Event (plain English description) | Event (set description) |
|---|---|
| All tosses come up heads | $\{HHH\}$ |
| All tosses come up the same | $\{HHH, TTT\}$ |
| Toss one head and two tails | $\{HTT, THT, TTH\}$ |
| First toss comes up heads | $\{HTT, HTH, HHT, HHH\}$ |
| The tosses don't come up all the same | $\{HHT, HTH, HTT, THH, THT, TTH\}$ |
| Toss more heads than tails | $\{HHT, HTH, THH, HHH\}$ |
| Toss an equal number of heads and tails | $\{\ \}$ (the *impossible* event) |
| Toss three or fewer heads | $S$ (the certain event) |

■ **TABLE 16-2** Same of the 256 possible events when tossing a coin three times.

## 16.2 The Multiplication Rule, Permutations, and Combinations

In this section we will introduce some of the basic mathematical tools that are used to determine the number of outcomes in a sample space (or an event) without having to list them all and then tallying them one by one. We will use the word **counting** to describe the process of determining the number of elements in a set, and our counting methods will be considerably more sophisticated than the kind of counting we learned in kindergarten.

You are probably already familiar with the first rule of counting—*the multiplication rule*.

- **Multiplication Rule.** If there are $m$ different ways to do $X$ and $n$ different ways to do $Y$, then $X$ and $Y$ together (and in that order) can be done in $m \times n$ different ways.

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 multiplication rule is a simple but powerful idea that we can best understand by looking at some examples.

**EXAMPLE 16.9** **ICE CREAM: EPISODE 1**

There is nothing sweeter than a great ice cream parlor, so we will use this theme in a few of our examples. Your local ice cream parlor offers 2 different choices of cones and 3 different flavors of ice cream (this is a boutique ice cream parlor!). Using the multiplication rule we can conclude that there are $2 \times 3 = 6$ different choices for a *single* (cone and one scoop of ice cream) order. The choices are illustrated in Figure 16-2.

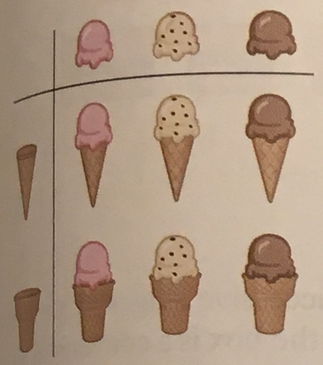


FIGURE 16-2 The scoop on the multiplication rule.

Figure 16-2 also clarifies the logic behind the multiplication rule: If we had $m$ different types of cones and $n$ different flavors of ice cream, then the corresponding arrangement would show the cone/flavor combinations laid out in $m$ rows and $n$ columns, illustrating the $m \times n$ possibilities.

**EXAMPLE 16.10** **ICE CREAM: EPISODE 2**

A bigger ice cream parlor around the corner offers 5 different choices of cones, 31 different flavors of ice cream, and 8 different choices of topping. The question is, If you are going to choose a cone and a single scoop of ice cream but then add a topping for good measure, how many orders are possible? The number of choices is easy to find using the multiplication rule twice: First, there are $5 \times 31 = 155$ different cone/flavor combinations, and each of these can be combined with one of the 8 toppings into a grand total of $155 \times 8 = 1240$ different cone/flavor/topping combinations. The implications—as well as the calories—are staggering!

**EXAMPLE 16.11** **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.)

Let's consider now a similar wardrobe question but with a few twists. 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?

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. 16-3. The final count for the number of different outfits is $N = 3 \times 7 \times 27 \times 3 = 1701$.

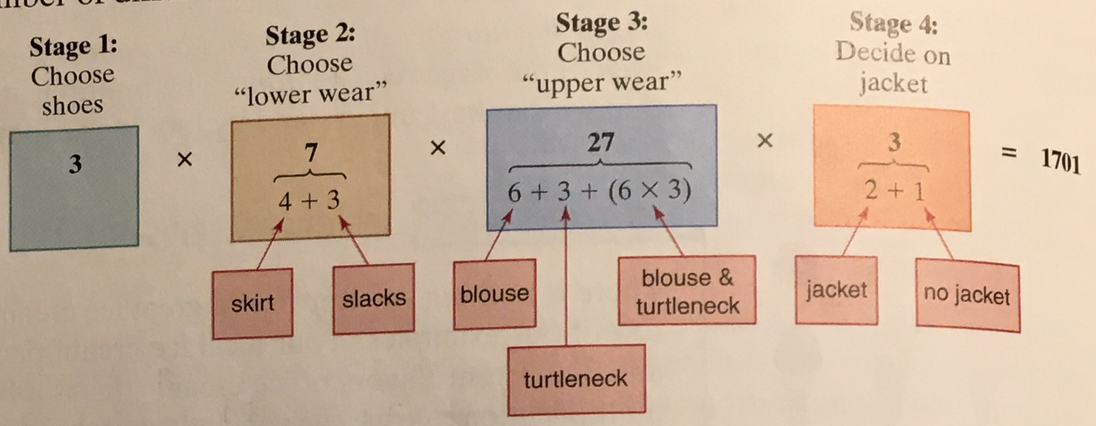



FIGURE 16-3

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 16.12 HORSE RACES AND TRIFECTAS: EPISODE 2

This example is a follow-up to Example 16.6. A trifecta ticket consists of three boxes labeled 1st, 2nd, and 3rd in which you can enter the numbers of the three horses you are betting will come in first, second, and third in the race. If there are 10 horses entered in the race, there are $10 \times 9 \times 8 = 720$ different outcomes in the sample space. Figure 16-4 illustrates the reasoning behind this count.

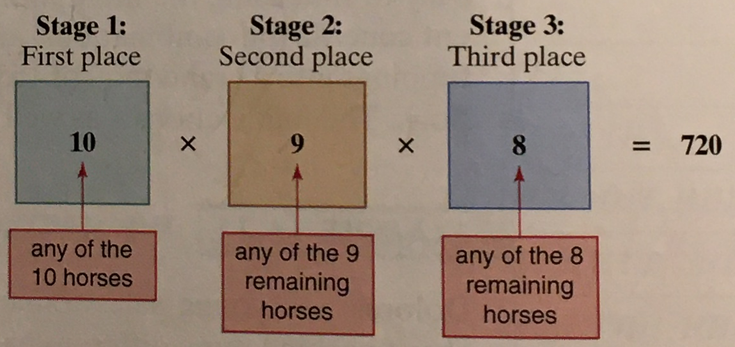



FIGURE 16-4

## Permutations and Combinations

Many counting problems boil down to counting the number of ways in which we can select 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*. A note of warning about the terminology: In Chapter 2 we defined the term *permutation* as a reordering or shuffling of a set of objects. We now stretch the meaning of the word.

## EXAMPLE 16.13 ICE CREAM: EPISODE 3

The ice cream parlor around the corner offers 31 different flavors of ice cream. One of their specials is the "true double"—two scoops of ice cream of *different* flavors served in a bowl and with chocolate sprinkles on top. Say you want to order a true double—how many different choices do you have?