READINGS

Probability

How often should we expect all three of the children in a family to be boys?

What are your chances of winning a sweepstakes contest or the door prize at a party?

How likely are you to land on "Boardwalk" in your next turn in a game of Monopoly?

If a baseball player averages 3 hits in 10 times at bat, what are his chances of getting 4 consecutive hits?

Have you ever wondered about the answers to such questions? Scientists often have occasion to study questions of this general type (though much more complicated than the four given above). In seeking the answers, they apply the mathematical theory of *probability*.

Probability is a mathematician's way of describing the *likelihood* that a certain event will take place. It is used to predict the outcome of an experiment when this outcome is governed by the laws of chance. Probability theory enables us to determine probable characteristics of a sample drawn from a population whose characteristics are known. It also provides the basis for part of the related science of *statistics*. Statistical inference is an extension of probability theory. In it, the outcome of an experiment is used to estimate the conditions governing the experiment. We would be drawing a statistical inference if we were to make an educated guess about some population by examining the characteristics of a sample taken from that population.

Let us now consider some examples. Suppose we know that there are 400 boys and 200 girls in a certain school and we conduct the experiment of choosing 1 student from this school at random. In a random selection, each of the 600 students would have exactly the same chance of being chosen. In this experiment, we would say that the probability of choosing a boy is 400 ÷ 600, or $^{2}/_{3}$, and the probability of choosing a girl is 200 ÷ 600, or $^{1}/_{3}$. The numbers $^{2}/_{3}$ and $^{1}/_{3}$ indicate that on the average we should expect to choose a boy "two times out of three" and a girl "one time out of three."

If we repeated this experiment 15 times, our best estimate would be that the sample chosen should contain $^{2}/_{3}$ boys and $^{1}/_{3}$ girls. This is an example of how probability is used to predict the outcome of an experiment. The probable composition of the sample was determined from the known composition of the population from which it was chosen.

Suppose that in a second school we do not know the proportion of boys and girls but know only that the total number of students is 600. To estimate the ratio of boys to girls, we might conduct an experiment consisting of selecting 20 students at random. If the sample chosen in this experiment contains 10 boys and 10 girls, our best estimate would be that there are equal numbers of boys and girls in the school—that there are 300 boys and 300 girls. In this case, the composition of the sample permits us to make a statistical inference about the composition of the population from which it was drawn.

As you see from these examples, probability and statistical inference are very closely related. The latter may be thought of as an application of probability in which the reasoning process is reversed. Statistical inference is only one aspect of the field of statistics, which also involves the collection, presentation, and analysis of data.

The examples just given illustrate the basic idea of probability and suggest how it may be applied, but they also may be misleading. They were oversimplified in order to avoid some of the more difficult problems that arise in the study of chance phenomena. In the first example, if we repeated the experiment a very large number of times, we might expect to choose a boy at random instead of a girl just about two times out of three. But the chances of selecting exactly 10 boys in a given sample of 15 are really quite small. This would happen only about two times in any ten trials. Most of the samples of 15 would be close to the theoretical ratio of $^{2}/_{3}$ (11 boys and 4 girls or 9 boys and 6 girls, for example), but some would not be close at all. We might even choose a random sample of 15 made up entirely of boys. In other words, the "most probable" outcome may not be very probable at all.

In the second example, the fact that our sample contained an equal number of boys and girls could lead us to a very poor estimate of the proportion of boys and girls in the school. We might even have drawn such a sample from the first school, in which there were twice as many boys as girls. In that case, we would have estimated the probability of choosing a boy as $^{1}/_{2}$ when it was actually $^{2}/_{3}$.

Many interesting and useful problems in probability are complicated because there are so many *possible* outcomes. If the number of possible outcomes is infinite, we must use calculus in our analysis. In this short introduction to probability theory, we shall deal with simple experiments producing a relatively small number of outcomes. However, the basic principles of probability are the same in simple problems as in more complicated ones. As you read, try to understand these principles, so that you can apply them to more complex situations. It will help you to grasp the subject if you work out all the situations given in the text as well as the special problems.

Probability and Games of Chance

Games of chance involving coins, cards, and dice provide us with simple experiments producing a small number of outcomes. The use of such experiments to illustrate the principles of probability is historically appropriate. Historians tell us that a 17th-century French gambler, the Chevalier de Méré, was interested in the odds involved in a game of chance played with dice. He decided to get in touch with a famous mathematician and scientist, Blaise Pascal, so that the latter might help him with his calculations. Pascal became intrigued with the questions that arose in his study of de Méré's problem. He began a correspondence with other mathematicians concerning the matter, and this led to the development of probability theory.

Consider the simple experiment of tossing coins. There are two possible outcomes when we toss a single coin: heads or tails. (We would ignore any toss in which the coin came to rest on its edge.) If the coin is in fair condition and if we toss it vigorously, it seems reasonable to say that heads and tails are equally likely outcomes. A mathematician would say that the probability of heads is $^{1}/_{2}$—that is, that the coin would land heads one time in two on the average.

Suppose we complicate this experiment just a little by tossing the coin twice or, what comes to about the same thing, tossing two coins. We can see that three things might happen in this experiment. We could get two heads, or one head and one tail, or two tails. The diagram in Figure 1 illustrates the possible outcomes in this experiment.

If we examine the diagram carefully, letting H stand for "heads" and T for "tails," we see that there are really four individual outcomes, or elementary events, possible—HH, HT, TH, and TT—when we toss two coins. Note that HT is not considered the same thing as TH. We might, for example, use a nickel and a dime as our two coins.

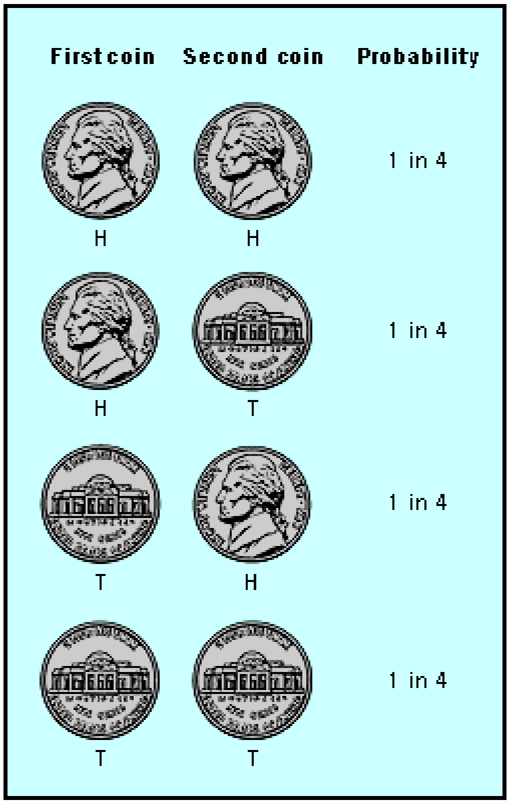

Figure 1. Possible outcomes in the tossing of two coins. An analysis of the likelihood of each of the four possible outcomes appears in the text.

Since each of the four individual outcomes is equally likely to occur, we would expect to obtain HH one time in four, or, as a mathematician would indicate it, $P(\text{HH}) = 1/4$. Similarly, we would say $P(\text{HT}) = 1/4$; $P(\text{TH}) = 1/4$; $P(\text{TT}) = 1/4$. Because what happens to the first coin has no effect upon what happens to the second coin, we say that the two tosses are *independent*. When two events are independent, we can use their individual probabilities to compute the probability that both will happen in a single trial of an experiment. In this case, we could have used the probabilities associated with tossing a single coin (that is, $P(\text{H}) = 1/2$; $P(\text{T}) = 1/2$) in order to compute the probabilities associated with tossing two coins. For example, $P(\text{HT}) = P(\text{H}) \times P(\text{T}) = 1/2 \times 1/2 = 1/4$; $P(\text{HH}) = P(\text{H}) \times P(\text{H}) = 1/2 \times 1/2 = 1/4$. Since the event "one head and one tail" could occur in two ways, either HT or TH, we would expect to obtain "one head and one tail" two times in four on the average and would assign probability to this event.

We could also use the individual probabilities assigned to the outcomes HT and TH to compute the probability that *either* HT *or* TH will occur. When two events cannot both occur in a single trial of an experiment, the probability that one or the other will occur is the sum of their individual probabilities. Thus $P(\text{HT or TH}) = \text{P}(\text{HT}) + P(\text{TH}) = 1/4 + 1/4 = 1/2$. In these examples, we obtain the probability of an event either by counting outcomes of an experiment or by using previously determined probabilities.

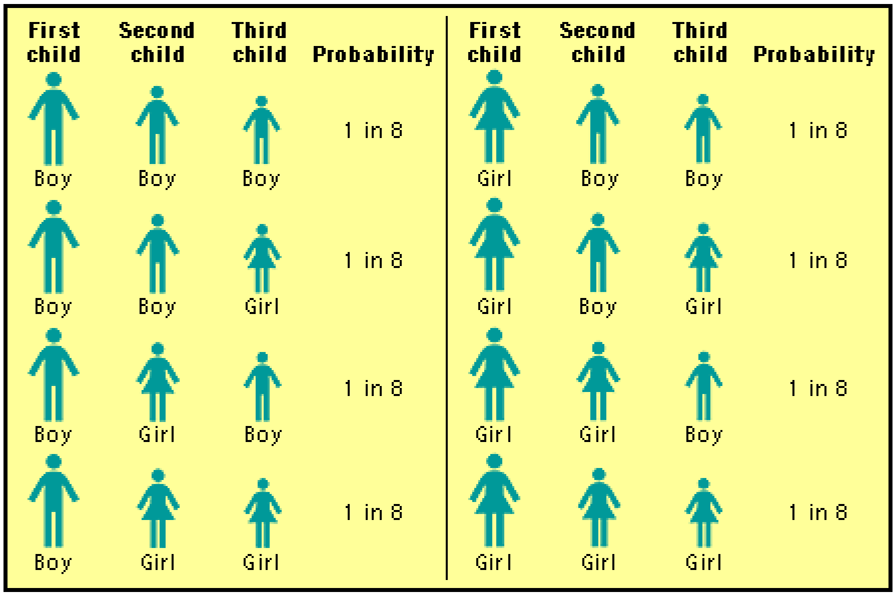

Figure 2. The diagram shows that the probability of having either three girls or three boys in a family of three children is 1 in 8. The probability of the first child being a boy is 1 in 2; the probability of the second being a boy is 1 in 2; and the probability of the third being a boy is 1 in 2.

We can apply the same principles in other experiments. Consider the question "How often should we expect all three of the children in a family to be boys?" The diagram in Figure 2 illustrates the possible outcomes. If we assumed that equal numbers of boys and girls are born (this is not quite true), we would say that the probability of a boy, P(B), is $^1/_2$ and that the probability of a girl, P(G), is also $^1/_2$. The possible outcomes for this experiment would be equal in number to the outcomes for tossing three coins—BBB, BBG, BGB, BGG, GBB, GBG, GGB, and GGG. Since BBB occurs in only one of the eight possible outcomes, we would say P(BBB) = $^1/_8$. We could also have used the rule for computing the probability that all of several independent events will occur: P(BBB) = P(B) × P(B) × P(B) = $^1/_2$ × $^1/_2$t × $^1/_2$ = $^1/_8$. Either way, we would conclude that we should expect all three children in a family to be boys about one time in eight.

Sample Spaces

To determine the probability of a particular outcome of an experiment, we must be able to identify all the possible outcomes. Mathematicians call the set of possible outcomes of an experiment a *sample space* for the experiment. In the simple examples, we shall usually find it convenient to list the sample spaces.

When we considered the experiment of tossing a single coin, our sample space consisted of the two possible outcomes H and T (heads and tails). When we extended the experiment to the tossing of two coins (or one coin twice), we used a sample space of HH, HT, TH, TT. In the experiment of choosing a single student from a known school population, the sample space involved was the set of all students in the school. Since we were concerned only with whether we chose a boy or girl and not with which boy or girl, we could have used a sample space of only two elements—boy and girl. In this case, B could be used to stand for the set of boys and G for the set of girls. Each individual outcome of a sample space is called a *point* or an *elementary event*.

Although a single experiment will produce only one set of individual outcomes, we may be able to use any one of several different sample spaces, depending on what we are investigating. For example, consider the experiment of drawing one card from a well-shuffled deck of cards. If we are concerned with the individual card drawn, we could use a sample space consisting of 52 elements—every single card in the deck. In the same experiment, we might be concerned only with the face value of the card shown. In that case, our sample space would consist of only 13 elements (2, 3, 4, 5, 6, 7, 8, 9, 10, jack, queen, king, ace). Or we might be interested only in the suit drawn. In that case, we would use the sample space of clubs, diamonds, hearts, spades. If we were concerned only with color, we could use a sample space of only two elements (red, black). In most experiments, we use the sample space identifying the characteristics on which we wish to concentrate in the experiment. We can use the sample space of elementary events to build up other sample spaces by collecting the elementary events with like characteristics.

You have probably played board games, such as Monopoly and Parcheesi, in which your moves were determined by the throw of a die or a pair of dice. A sample space for the experiment of tossing a single die would be the set {1, 2, 3, 4, 5, 6}. It is rather more difficult to generate a sample space for the tossing of a pair of dice. To help us keep things straight, let us suppose that one die is red and the other one green. An outcome of 4 on the red die and 3 on the green die would be different from an outcome of 3 on the red die and 4 on the green die. If we agreed to write the outcome on the red die first and the outcome on the green die second, we could show this difference by setting down the pairs (4,3) and (3,4).

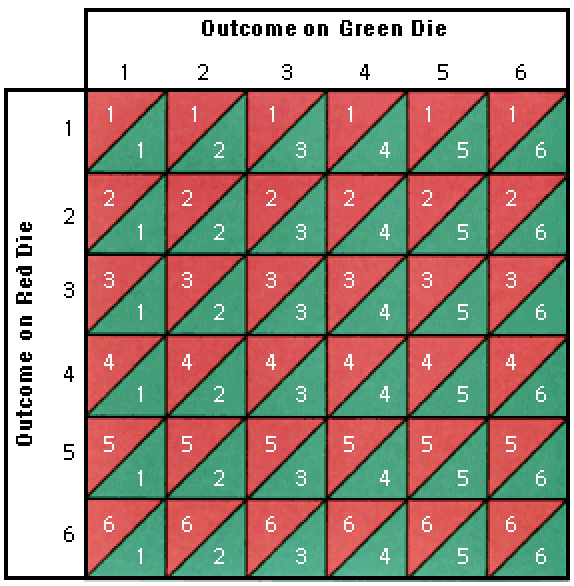

Figure 3. Here is a sample space for the tossing of a pair of dice. It is assumed that one of the dice is red and the other green; also that the outcome of the red die is given first in each of the pairs shown. Note that in the text the two outcomes, separated by a comma, are given in parentheses. Thus we would refer to (2,5), (4,3), (6,1), and so on.

Figure 3 shows the sample space for the experiment of tossing a pair of dice. Suppose that you tossed the dice and that the red die came up 4 and the green die came up 3. Moving horizontally from 4, at the extreme left of the table, until we came to the column headed by 3, we would find the pair (4,3). We shall have occasion to refer to this sample space several times in the course of this article.

See if you can list sample spaces for the following experiments:

1. Toss a coin and a die. [Hint: two points in this sample space could be (H,3) and (T,2).]
2. Toss a nickel, dime, and quarter. If the nickel comes up heads, the dime tails, and the quarter heads, we can indicate this by (H,T,H). You should find 8 points for this sample space.

Probability of an Event

An *event* is by definition any subset of a sample space. In other words, if each member of set A is also a member of set B, we say that A is a subset of B. If set A = {1, 2, 3} and set B = {1, 2, 3, 4, 5}, then set A is a subset of set B. Thus an event is a set of individual outcomes of an experiment. An elementary event, as we have indicated, is a single individual outcome. In order to assign probabilities to an event— that is, to a set of individual outcomes—we must first be able to assign probabilities to individual outcomes of the experiment.

If we consider, for example, the experiment of drawing a single card from a deck of cards, there are 52 elementary events or individual outcomes possible. If we make a random draw, each card has exactly the same chance of being drawn. Thus to each of the 52 possible outcomes we would assign the same probability: $^{1}/_{52}$. Note that the sum of the probabilities assigned to the elementary events is 1. The probability of an event that is certain is also 1.

The same basic principle may be used to answer the question "What are your chances of winning a sweepstakes contest or the door prize at a party?" In each case, if the total number of tickets is *n* and you hold only 1 ticket, then you would have only 1 chance in *n* of winning. Thus the probability of your winning would be $^{1}/_{n}$.

In the experiment of drawing a card from a deck, we may be concerned only with whether the card is an ace. Then the event with which we are concerned consists of four elementary events: the ace of clubs, the ace of diamonds, the ace of hearts, and the ace of spades. We still would have to draw at random from the entire deck of cards, numbering 52. Since 4 out of the 52 cards belong to the event just described (4 aces), we would say that the probability that the outcome is an ace is $^{4}/_{52}$. We could use the symbols $P(\text{ace}) = {}^{4}/_{52}$ to represent this statement.

Suppose that in the preceding experiment, we were concerned only with the color of the card chosen. Do you see why $P(\text{red}) = {}^{26}/_{52}$? We can use the probabilities assigned to elementary events to assign probabilities to the points in other sample spaces. For example, if we were using the sample space clubs, diamonds, hearts, spades, we could assign the probability of $^{1}/_{4}$ to each point in the sample space since $P(\text{clubs}) = {}^{13}/_{52} = {}^{1}/_{4}$.

Rule 1. If an experiment can result in *n* different but equally likely outcomes and if *m* of these outcomes correspond to event X, then the probability of the event is $P(X) = {}^{m}/_{n}$.

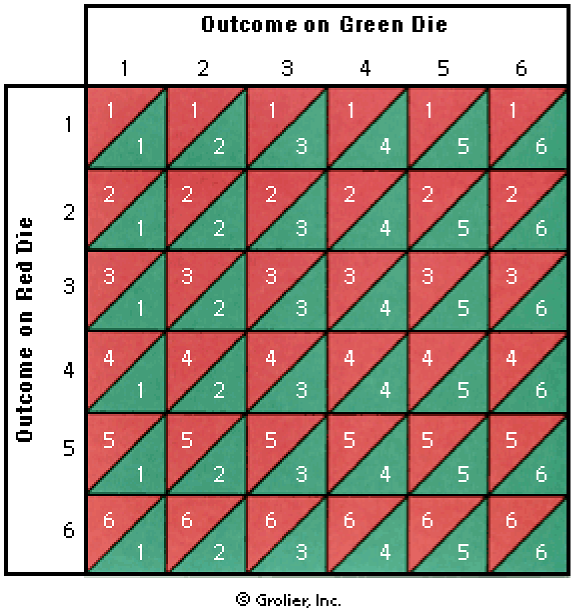

Figure 3. Here is a sample space for the tossing of a pair of dice. It is assumed that one of the dice is red and the other green; also that the outcome of the red die is given first in each of the pairs shown. Note that in the text the two outcomes, separated by a comma, are given in parentheses. Thus we would refer to (2,5), (4,3), (6,1), and so on.

Applying this rule, let us find the probabilities of some events in the experiment of tossing a pair of dice. Figure 3 lists 36 possible outcomes for this experiment. Thus we could assign a probability of $^{1}/_{36}$ to each of the elementary events. Now consider the event "The sum of the numbers shown is 7." If we let r stand for the number on the red die and g for the number on the green die, we can represent the event as $r + g = 7$. How many of the elementary events correspond to this amount? In other words, how many of the pairs in the table add up to 7? Consult the table to find the answer. You will see that $P(r + g = 7) = {}^{6}/_{36}$. If we considered the event "The same number appears on both dice," we could call the event "$r = g$" or, as is common in many games played with dice, "double." There are six such pairs: (1,1), (2,2), (3,3), (4,4), (5,5), and (6,6). Thus $P(r + g) = P(\text{double}) = {}^{6}/_{36}$. What is $P(r + g > 7)$? (The symbol > stands for "is greater than.") Are you able to find 15 points in the sample space that correspond to this event? You should be able to, for $P(r + g > 7) = {}^{15}/_{36}$.

These examples should suggest a way to answer the question "How likely are you to land on 'Boardwalk' in your next turn in a game of Monopoly?" Suppose, for example, that you are located on "Pennsylvania Avenue," which is 5 spaces from "Boardwalk." To determine the probability of your landing on "Boardwalk," carry out the experiment of tossing two dice. You would land on "Boardwalk" if the numbers shown on the dice totaled 5. What is $P(r + g = 5)$?

Try the following problems:

1. If a box contains 4 red marbles and 5 white marbles, what is the probability of drawing a red marble on the first try?
2. What is the probability that you will draw a face card (jack, queen, king) from a deck of cards?
3. What is the probability that you will get a 5 when you toss a single die?

4. What is the probability that you will not get a 5 when you toss a single die?

Complementary Events

There is a relationship that frequently simplifies the computing of a probability. Consider the example pictured in Figure 4. There are 10 points in the sample space S. The event A contains 4 of these points and the remaining 6 points are not in A. The set of points in sample space S that

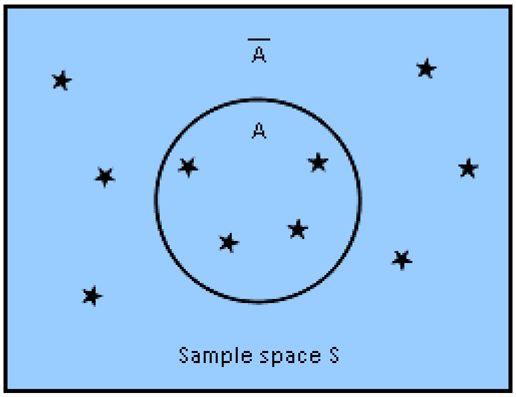

Figure 4. Subsets A and A-macron are complementary events in sample space S.

are not in a given set A is called the *complement* of A and is usually indicated by the symbol $(\overline{A})$.

If the elementary events in the sample space S are equally likely, then $P(A) = {}^{4}/_{10}$ and $P(\overline{A}) = \frac{6}{10}$.

In a sample space of n equally likely events, if an event A occurs in m of the outcomes, then the event $(\overline{A})$ will occur in $n - m$ outcomes. Thus if $P(A) = {}^{m}/_{n}$, then $P(\overline{A}) = \frac{(n-m)}{n}$. We obtain this result by dividing both the numerator and the denominator by n, as follows:

$$\frac{\frac{n}{n} - \frac{m}{n}}{\frac{n}{n}} = \frac{1 - \frac{m}{n}}{1} = 1 - \frac{m}{n}$$

We now can give the following general rule:

Rule 2.

$$P(\overline{A}) = 1 - P(A)$$

Sometimes it is easier to compute the probability of one of two complementary events than it is to compute the probability of the other. In such cases, we compute the easier probability. We then use the relationship indicated in Rule 2 to derive the probability of the complementary event. For example, in some games played with dice, there is either a premium or a penalty associated with throwing doubles. We could compute the probability of not throwing a double by counting the sample points in Figure 3 that are not doubles. (There are 30 of them.) Or we could compute the probability of throwing a double ($r = g$) and use the relationship in Rule 2 to derive the probability of not throwing a double ($r \neq g$). (The symbol ≠ stands for "is not equal to.")

Probability of A or B

A situation that often arises is the finding of the probability of an event that might be expressed as "either event A or event B." By the event (A or B) we mean the set of outcomes that correspond to either event A or event B, or possibly both A and B.

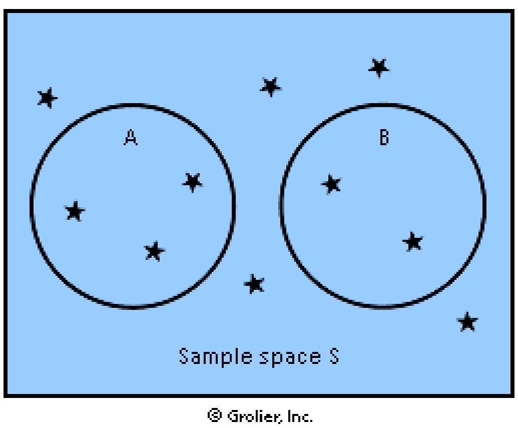

Figure 5. Subsets A and B are mutually exclusive events in sample space S.

Consider the example shown in Figure 5. There are 10 points in the sample space S. In this sample space, $P(A) = 3/10$ and $P(B) = 2/10$. Since the event (A or B) contains the 3 elements of A and the 2 elements of B, then $P(A \text{ or } B) = P(A) + P(B) = 5/10$. In this instance, there are no elements that are in both A and B; so A and B here are *mutually exclusive* events, meaning that they cannot both occur in a single trial of an experiment. When events A and B are mutually exclusive, we can find the probability of (A or B) simply by adding the probability of A to the probability of B as follows:

Rule 3. If events A and B are mutually exclusive, then $P(A \text{ or } B) = P(A) + P(B)$.

For example, suppose we were required to find the probability of throwing either a 7 or a double in a toss of two dice. Since 7 is an odd number, it is not possible for a single toss of two dice to produce both a 7 and a double. Thus the event ($r + g = 7$) and the event ($r = g$) are mutually exclusive. We have already seen that $P(r + g = 7) = 6/36$ and that $P(r = g) = 6/36$. Hence, applying

Rule 3, $P(r + g = 7 \text{ or } r = g) = {}^{6}/_{36} + {}^{6}/_{36} = {}^{12}/_{36}$. Can you verify this by counting the appropriate points in the sample space shown in Figure 3?

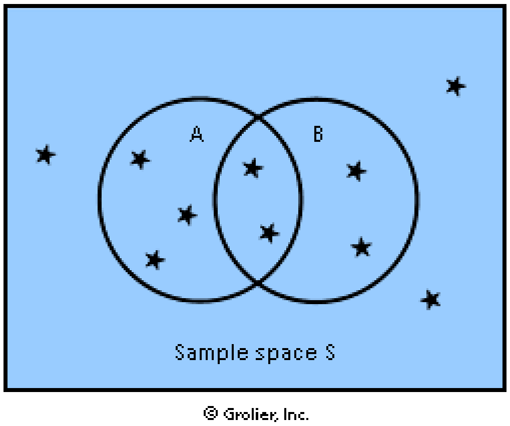

Figure 6. Subsets A and B are non-mutually exclusive events in sample space S.

When both events are not mutually exclusive —that is, when they *can* both occur in a single outcome—we cannot use the addition principle of Rule 3. Consider the situation presented in Figure 6. In this example, the sample space S consists of 10 equally likely elementary events; 5 outcomes correspond to event A, and 4 outcomes to event B. Hence $P(A) = {}^{5}/_{10}$ and $P(B) = {}^{4}/_{10}$. If we were to apply Rule 3 here, we would have $P(A \text{ or } B) = P(A) + P(B) = {}^{5}/_{10} + {}^{4}/_{10} = {}^{9}/_{10}$. This answer would be incorrect. The reason is that events A and B are not mutually exclusive, since 2 outcomes belong to both A and B. To obtain a correct answer in this case for $P(A \text{ or } B)$, we need to apply the following rule:

Rule 4. If events A and B are not mutually exclusive, then $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$.

Therefore, for Figure 6, $P(A \text{ or } B) = {}^{5}/_{10} + {}^{4}/_{10} - {}^{2}/_{10} = {}^{7}/_{10}$.

Let us consider another example in which this rule could be used. Suppose we wanted to find the probability of drawing either a face card or a spade from a deck of cards. $P(\text{spade}) = {}^{13}/_{52}$ and $P(\text{face card}) = {}^{12}/_{52}$. (Remember that each of the 4 suits has 3 face cards: jack, queen, and king.) There are 3 cards that are both face cards and spades. Therefore $P(\text{spade and face card}) = {}^{3}/_{52}$. Applying Rule 4, we would conclude that $P(\text{spade or face card}) = {}^{13}/_{52} + {}^{12}/_{52} - {}^{3}/_{52} = {}^{22}/_{52}$. Could you list the 22 cards that would belong to the event (spade or face card)?

Here are two rather simple problems for you to try:

1. What is the probability that you could throw either a 7 or an 11 on a single throw of 2 dice?
2. What is the probability that you would throw either a double or a total of more than 9 on a single toss of 2 dice?

Probability of A and B

When two events, A and B, are *independent*, we can compute the probability of the event (A and B) by using the following rule, involving the multiplication principle:

Rule 5. If events A and B are independent, then *P*(A and B) = *P* (A) × *P*(B).

Intuitively, we would expect two events to be independent when they have nothing to do with each other. Although this is usually a reliable rule of thumb, it must be used with care. For example, when two events are mutually exclusive, our first thought might be that they have nothing to do with each other. But when one of two mutually exclusive events occurs, the other cannot possibly occur. Thus the occurrence of one of these events certainly affects the probability that the other also occurs. Mutually exclusive events, therefore, are never independent of one another. For two events to be independent, the occurrence of one must not affect the probability that the other occurs at the same time.

Here is a problem involving two independent events, in which Rule 5 could be applied: Suppose that we toss two dice, one red and one green. What is the probability that the number 1 would come up on the red die and that the number 3 would come up on the green die? In this case, the outcome on the green die has nothing to do with the outcome on the red die. We therefore have two independent events. The probability of any one of the numbers 1, 2,3, 4, 5, and 6 coming up on the toss of a single die is $^1/_6$. The probability of the outcome (1,3) on a toss of two dice would be calculated as follows, in accordance with Rule 5: $P(1 \text{ and } 3) = P(1) \times P(3) = {}^1/_6 \times {}^1/_6 = {}^1/_{36}$.

By using the multiplication principle, we can provide an answer of sorts to the question "If a baseball player averages 3 hits in 10 times at bat, what are his chances of getting 4 consecutive hits?" If we assume that the times at bat are independent trials, and if we let H stand for each hit, we could compute the possibility of 4 consecutive hits in this way: $P(\text{HHHH}) = P(\text{H}) \times P(\text{H}) \times P(\text{H}) \times P(\text{H}) = {}^3/_{10} \times {}^3/_{10} \times {}^3/_{10} \times {}^3/_{10} = {}^{81}/_{10,000}$. We would expect the batter to get 4 hits in a row about 81 times in every 10,000 sequences of 4 times at bat. This probability would hold if we were right in assuming that the 4 times at bat would be independent trials. The assumption might not be a sound one in this case. For one thing, the batter would probably hit better against certain pitchers than against others. Also, getting a hit or not getting a hit might have an effect on his chances the next time at bat.

Consider the event (double and even) in the experiment of tossing two dice. By "double," of course, we mean that the same number would come up on both dice; by "even," that the sum of the two numbers would be an even number. In this we would have $P(\text{double}) = {}^6/_{36}$ and $P(\text{even}) = {}^{18}/_{36}$. If we applied the multiplication principle for independent events, as stated in Rule 5, we would have $P(\text{double and even}) = {}^6/_{36} \times {}^{18}/_{36} = {}^{108}/_{1,296} = {}^3/_{36}$. But if we examine Figure 3, we will find 6 points that correspond to the event (double and even). Therefore the true value of *P*(double and even) is $^6/_{36}$ and not $^3/_{36}$, as our calculation had seemed to indicate. The reason is that in this case the two events are not independent. If we throw a double, we are certain to throw an even number. This means that if we have thrown a double, the probability that we have also thrown an even number is 1. Again, if we know that we have thrown an even number, the probability that we have also thrown a double is $^1/_3$. This is determined by counting the number of ways an even number can occur (18 in all and the number of these occurrences that are doubles (6).

In order to calculate *P*(A and B) when the events are not independent, the multiplication principle for independent events (Rule 5) can be generalized as follows:

Rule 6. If events A and B are not independent, then *P*(A and B) = *P*(A) × *P*(B/A).

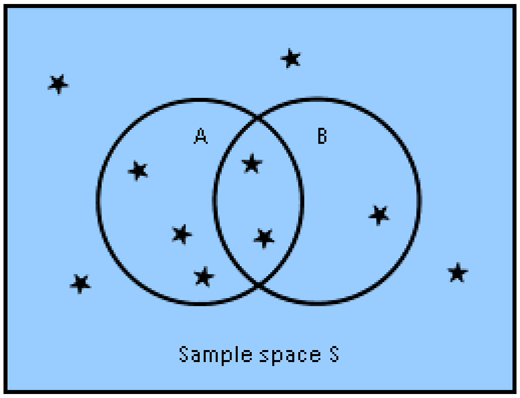

Figure 7. Subsets A and B are nonindependent events in sample space S.

The term "B/A" would be read as "B given A." When we know that the event A occurs, we can compute *P*(B/A) by thinking of the event A as a reduced sample space. Consider Figure 7. The sample space S consists of 10 equally likely events, of which 5 are in A, 3 are in B, and 2 are in both A and B. Given that A occurs, then any outcome obtained is one of the 5 in A. If we think of these 5 points as a new sample space, then 2 of the equally likely outcomes belong to the event B. Thus *P*(B/A) = $^{2}/_{5}$. Applying Rule 6, we can compute *P*(A and B) = *P*(A) × *P*(B/A) = $^{5}/_{10}$ × $^{2}/_{5}$ = $^{10}/_{50}$ = $^{2}/_{10}$. This, as you can see, agrees with what we would obtain by a direct count of the points that belong to both A and B in Figure 7.

If events A and B are independent, then the probability that B occurs will not be affected by the fact that A occurs. Thus for independent events, *P*(B/A) = *P*(B). Suppose we were to select 1 card from a deck of 52 cards. What is the probability that the card selected will be both a red card (R) and a face card (F)? Since half the cards are red, we know that *P*(R) = $^{26}/_{52}$. Of the 26 red cards, 6 are face cards. Hence *P*(F/R) = $^{6}/_{26}$. Since *P*(F) = $^{12}/_{52}$ = $^{6}/_{26}$ and *P*(F/R) is also $^{6}/_{26}$, the events F and R are independent and *P*(R and F) = *P*(R) × *P*(F). Thus the multiplication rule for independent events (Rule 5) is actually just a special case of the more general Rule 6: *P*(A and B) = *P*(A) × *P*(B/A).

Here are some problems involving the probability of event (A and B). (Assume that cards are drawn from a regular 52-card deck.)

1. A coin is tossed and a die is thrown. What is the probability of obtaining heads and a 3?
2. A card is drawn and a die is tossed. What is the probability of getting two 6's?
3. A card is drawn. What is the probability that it is neither a face card nor a red card?

Degree of Confidence

In this brief introduction to probability, we have touched on only a few of the basic principles involved. The scope of probability theory is much broader than our discussion would suggest. For example, we simplified matters a good deal by dealing only with situations in which the individual outcomes of an experiment were equally likely. In many situations this is not the case. Calculations then become more complex, and new dimensions are added to the study of probability.

In our simplified presentation, we merely pointed out that predictions based on probability are uncertain. Scientists accept this but want to know *how* uncertain. They recognize that in trying to predict the outcome of an experiment subject to the laws of chance, they may often be wrong. They need to know how often and how wrong. They want to establish the degree of confidence they can place in their predictions. Probability theory provides the basis for establishing this degree of confidence.

Many Uses

If you go at all deeply into the study of mathematics, you will find occasion to work with some of the more complicated and intriguing aspects of probability theory. You will also obtain a clearer idea of the many ways in which this theory serves humans: it enables scientists to fix a limit within which the deviations from a given physical law must fall if these deviations are not to count against the law. It has been used to calculate the positions and velocities of electrons orbiting around the nuclei of atoms. The fluctuations in density of a given volume of gas have been analyzed by applying probability theory. It has played an important part in genetics; among other things, it has made it possible to calculate the percentage of individuals with like and unlike traits in successive generations. Manufacturers use probability theory to predict the quality of items coming off mass-production lines. Insurance experts make extensive use of the theory; it enables them, for example, to calculate life expectancies so that they may set appropriate life-insurance rates. Finally, because of probability theory, computers can be programmed to predict the outcome of elections on the basis of comparatively few returns and with what is generally a surprising degree of accuracy.

F. Joe Crosswhite

It's always a pleasant surprise to find that someone you know has the same birthday as you do. But how much of a coincidence is it? Put another way, what is the minimum number of people you need to have in a group to make it more likely than not that two of the members will share a birthday (month and day, but not year)? It turns out that the minimum number is unexpectedly small—23, in fact. The probability can be computed as follows: In a "group" of one person, there is no possibility of coinciding birthdays, so the probability of no match is 366/366, or 1. Now add a second person. For that person not to share a birthday with the first person, he or she must be born on one of the other 365 days; that is, the probability of the dates not matching is 365/366. To get the probability that the birthdays *will* match, simply subtract the no-match probability figure from 1. Therefore, if there are two people, the probability that the birthdays will match is 1 - 0.9973, or 0.0027. If a third person joins the group, the no-match probability is 364/366, or 0.995, and the probability of a match becomes 0.005. The process continues as you keep adding people, until finally, at the 23rd person, the no-match probability becomes 0.4927, which makes the chance of a match 0.5073, or slightly better than 50 percent.