



CHAPTER 4

Introduction to Probability

CONTENTS

STATISTICS IN PRACTICE: ROHM AND HASS COMPANY

4.1 EXPERIMENTS, COUNTING RULES, AND ASSIGNING PROBABILITIES

Counting Rules, Combinations, and Permutations
Assigning Probabilities
Probabilities for the KP&L Project

4.2 EVENTS AND THEIR PROBABILITIES

4.3 SOME BASIC RELATIONSHIPS OF PROBABILITY

Complement of an Event
Addition Law

4.4 CONDITIONAL PROBABILITY

Independent Events
Multiplication Law

4.5 BAYES' THEOREM

Tabular Approach

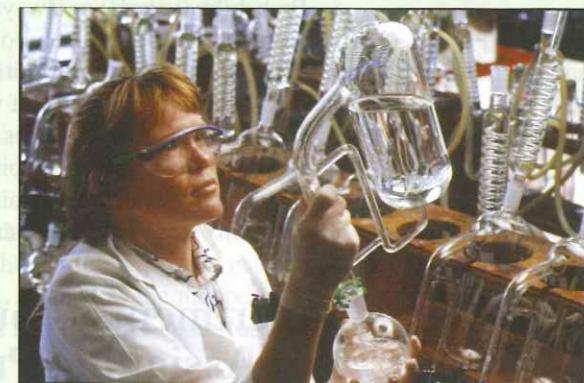
STATISTICS *in* PRACTICE

ROHM AND HASS COMPANY* PHILADELPHIA, PENNSYLVANIA

Rohm and Hass is the world's leading producer of specialty materials, including electronic materials, polymers for paints, and personal care items. Company products enable the creation of leading-edge consumer goods in markets such as pharmaceuticals, retail food, building supplies, communication equipment, and household products. The company has a workforce of more than 17,000 and annual sales of \$8 billion. A network of more than 100 manufacturing, technical research, and customer service sites provide Rohm and Hass products and service in 27 countries worldwide.

In the area of specialty chemical products, the company offers a variety of chemicals designed to meet the unique specifications of its customers. For one particular customer, the company produced an expensive catalyst used in the customer's chemical processing operation. Some, but not all, of the shipments from the company met the customer's specifications for the product. The contract called for the customer to test each shipment after receiving it and determine whether the catalyst would perform the desired function. Shipments that did not pass the customer's test would be returned. Over time, experience showed that the customer was accepting 60% of the shipments, but returning 40% of the shipments. Neither the customer nor the company was pleased with this level of service.

The company explored the possibility of duplicating the customer's test prior to shipment. However, the high cost of the special testing equipment that was required made this alternative infeasible. Company chemists working on the problem proposed a different but relatively low-cost test that could be conducted prior to shipment to the customer. The company believed that the new test would provide an indication of whether the catalyst would pass the customer's



A new test prior to shipment improved customer service. © Keith Wood/Stone.

more sophisticated test. The probability question was: What is the probability that the catalyst would pass the customer's test given that it passed the new test prior to shipment?

A sample of the catalyst was produced and subjected to the new company test. Only samples of the catalyst that passed the new test were sent to the customer. Probability analysis of the data indicated that if the catalyst passed the new test prior to shipment, there was a .909 probability that the catalyst would pass the customer's test. Or, if the catalyst passed the new test prior to shipment, there was only a .091 probability that it would fail the customer's test and be returned. The probability analysis provided supporting evidence for the implementation of the testing procedure prior to shipment. This new test resulted in an immediate improvement in customer service and a substantial reduction in shipping and handling costs for the returned shipments.

The probability of a shipment being accepted by the customer given it had passed the new test is called a conditional probability. In this chapter, you will learn how to compute conditional and other probabilities that are helpful in decision making.

*The authors are indebted to Michael Haskell of the Rohm and Hass subsidiary Morton International for providing this statistics in practice.

Managers often base their decisions on an analysis of uncertainties such as the following:

1. What are the chances that sales will decrease if we increase prices?
2. What is the likelihood a new assembly method will increase productivity?
3. How likely is it that the project will be finished on time?
4. What is the chance that a new investment will be profitable?

Some of the earliest work on probability originated in a series of letters between Pierre de Fermat and Blaise Pascal in the 1650s.

Probability is a numerical measure of the likelihood that an event will occur. Thus, probabilities can be used as measures of the degree of uncertainty associated with the four events previously listed. If probabilities are available, we can determine the likelihood of each event occurring.

Probability values are always assigned on a scale from 0 to 1. A probability near zero indicates an event is unlikely to occur; a probability near 1 indicates an event is almost certain to occur. Other probabilities between 0 and 1 represent degrees of likelihood that an event will occur. For example, if we consider the event “rain tomorrow,” we understand that when the weather report indicates “a near-zero probability of rain,” it means almost no chance of rain. However, if a .90 probability of rain is reported, we know that rain is likely to occur. A .50 probability indicates that rain is just as likely to occur as not. Figure 4.1 depicts the view of probability as a numerical measure of the likelihood of an event occurring.

4.1

Experiments, Counting Rules, and Assigning Probabilities

In discussing probability, we define an **experiment** as a process that generates well-defined outcomes. On any single repetition of an experiment, one and only one of the possible experimental outcomes will occur. Several examples of experiments and their associated outcomes follow.

Experiment	Experimental Outcomes
Toss a coin	Head, tail
Select a part for inspection	Defective, nondefective
Conduct a sales call	Purchase, no purchase
Roll a die	1, 2, 3, 4, 5, 6
Play a football game	Win, lose, tie

By specifying all possible experimental outcomes, we identify the **sample space** for an experiment.

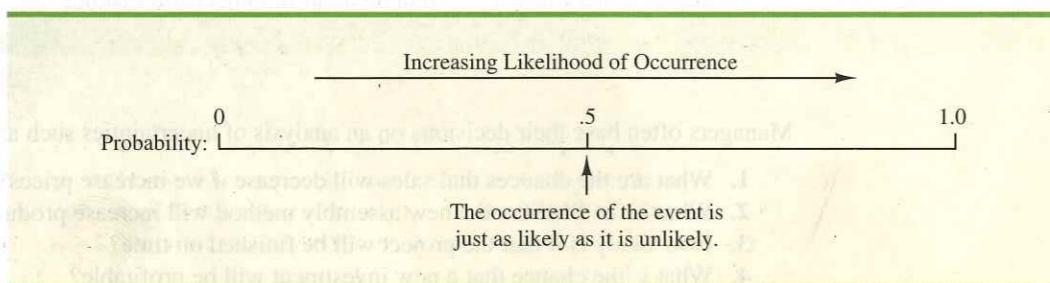
SAMPLE SPACE

The sample space for an experiment is the set of all experimental outcomes.

Experimental outcomes are also called sample points.

An experimental outcome is also called a **sample point** to identify it as an element of the sample space.

FIGURE 4.1 PROBABILITY AS A NUMERICAL MEASURE OF THE LIKELIHOOD OF AN EVENT OCCURRING



Consider the first experiment in the preceding table—tossing a coin. The upward face of the coin—a head or a tail—determines the experimental outcomes (sample points). If we let S denote the sample space, we can use the following notation to describe the sample space.

$$S = \{\text{Head, Tail}\}$$

The sample space for the second experiment in the table—selecting a part for inspection—can be described as follows:

$$S = \{\text{Defective, Nondefective}\}$$

Both of the experiments just described have two experimental outcomes (sample points). However, suppose we consider the fourth experiment listed in the table—rolling a die. The possible experimental outcomes, defined as the number of dots appearing on the upward face of the die, are the six points in the sample space for this experiment.

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

Counting Rules, Combinations, and Permutations

Being able to identify and count the experimental outcomes is a necessary step in assigning probabilities. We now discuss three useful counting rules.

Multiple-step experiments The first counting rule applies to multiple-step experiments. Consider the experiment of tossing two coins. Let the experimental outcomes be defined in terms of the pattern of heads and tails appearing on the upward faces of the two coins. How many experimental outcomes are possible for this experiment? The experiment of tossing two coins can be thought of as a two-step experiment in which step 1 is the tossing of the first coin and step 2 is the tossing of the second coin. If we use H to denote a head and T to denote a tail, (H, H) indicates the experimental outcome with a head on the first coin and a head on the second coin. Continuing this notation, we can describe the sample space (S) for this coin-tossing experiment as follows:

$$S = \{(H, H), (H, T), (T, H), (T, T)\}$$

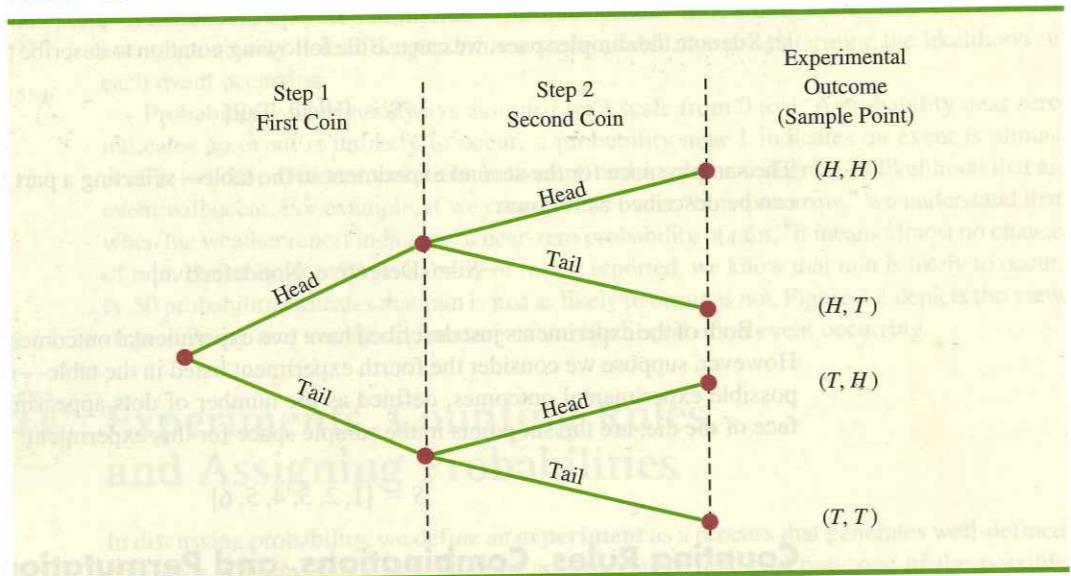
Thus, we see that four experimental outcomes are possible. In this case, we can easily list all of the experimental outcomes.

The counting rule for multiple-step experiments makes it possible to determine the number of experimental outcomes without listing them.

COUNTING RULE FOR MULTIPLE-STEP EXPERIMENTS

If an experiment can be described as a sequence of k steps with n_1 possible outcomes on the first step, n_2 possible outcomes on the second step, and so on, then the total number of experimental outcomes is given by $(n_1)(n_2) \dots (n_k)$.

If we view the experiment of tossing two coins as a sequence of first tossing one coin ($n_1 = 2$) and then tossing the other coin ($n_2 = 2$), the counting rule then shows us that $(2)(2) = 4$ distinct experimental outcomes are possible: $S = \{(H, H), (H, T), (T, H), (T, T)\}$. The number of experimental outcomes in an experiment involving tossing six coins is $(2)(2)(2)(2)(2)(2) = 64$.

FIGURE 4.2 TREE DIAGRAM FOR THE EXPERIMENT OF TOSSING TWO COINS

Without the tree diagram, one might think only three experimental outcomes are possible for two tosses of a coin: 0 heads, 1 head, and 2 heads.

A **tree diagram** is a graphical representation that helps in visualizing a multiple-step experiment. Figure 4.2 shows a tree diagram for the experiment of tossing two coins. The sequence of steps moves from left to right through the tree. Step 1 corresponds to tossing the first coin, and step 2 corresponds to tossing the second coin. For each step, the two possible outcomes are head or tail. Note that for each possible outcome at step 1 two branches correspond to the two possible outcomes at step 2. Each of the points on the right end of the tree corresponds to an experimental outcome. Each path through the tree from the left-most node to one of the nodes at the right side of the tree corresponds to a unique sequence of outcomes.

Let us now see how the counting rule for multiple-step experiments can be used in the analysis of a capacity expansion project for the Kentucky Power & Light Company (KP&L). KP&L is starting a project designed to increase the generating capacity of one of its plants in northern Kentucky. The project is divided into two sequential stages or steps: stage 1 (design) and stage 2 (construction). Even though each stage will be scheduled and controlled as closely as possible, management cannot predict beforehand the exact time required to complete each stage of the project. An analysis of similar construction projects revealed possible completion times for the design stage of 2, 3, or 4 months and possible completion times for the construction stage of 6, 7, or 8 months. In addition, because of the critical need for additional electrical power, management set a goal of 10 months for the completion of the entire project.

Because this project has three possible completion times for the design stage (step 1) and three possible completion times for the construction stage (step 2), the counting rule for multiple-step experiments can be applied here to determine a total of $(3)(3) = 9$ experimental outcomes. To describe the experimental outcomes, we use a two-number notation; for instance, (2, 6) indicates that the design stage is completed in 2 months and the construction stage is completed in 6 months. This experimental outcome results in a total of $2 + 6 = 8$ months to complete the entire project. Table 4.1 summarizes the nine experimental outcomes for the KP&L problem. The tree diagram in Figure 4.3 shows how the nine outcomes (sample points) occur.

The counting rule and tree diagram help the project manager identify the experimental outcomes and determine the possible project completion times. From the information in

TABLE 4.1 EXPERIMENTAL OUTCOMES (SAMPLE POINTS) FOR THE KP&L PROJECT

Completion Time (months)		Notation for Experimental Outcome	Total Project Completion Time (months)
Stage 1 Design	Stage 2 Construction		
2	6	(2, 6)	8
2	7	(2, 7)	9
2	8	(2, 8)	10
3	6	(3, 6)	9
3	7	(3, 7)	10
3	8	(3, 8)	11
4	6	(4, 6)	10
4	7	(4, 7)	11
4	8	(4, 8)	12

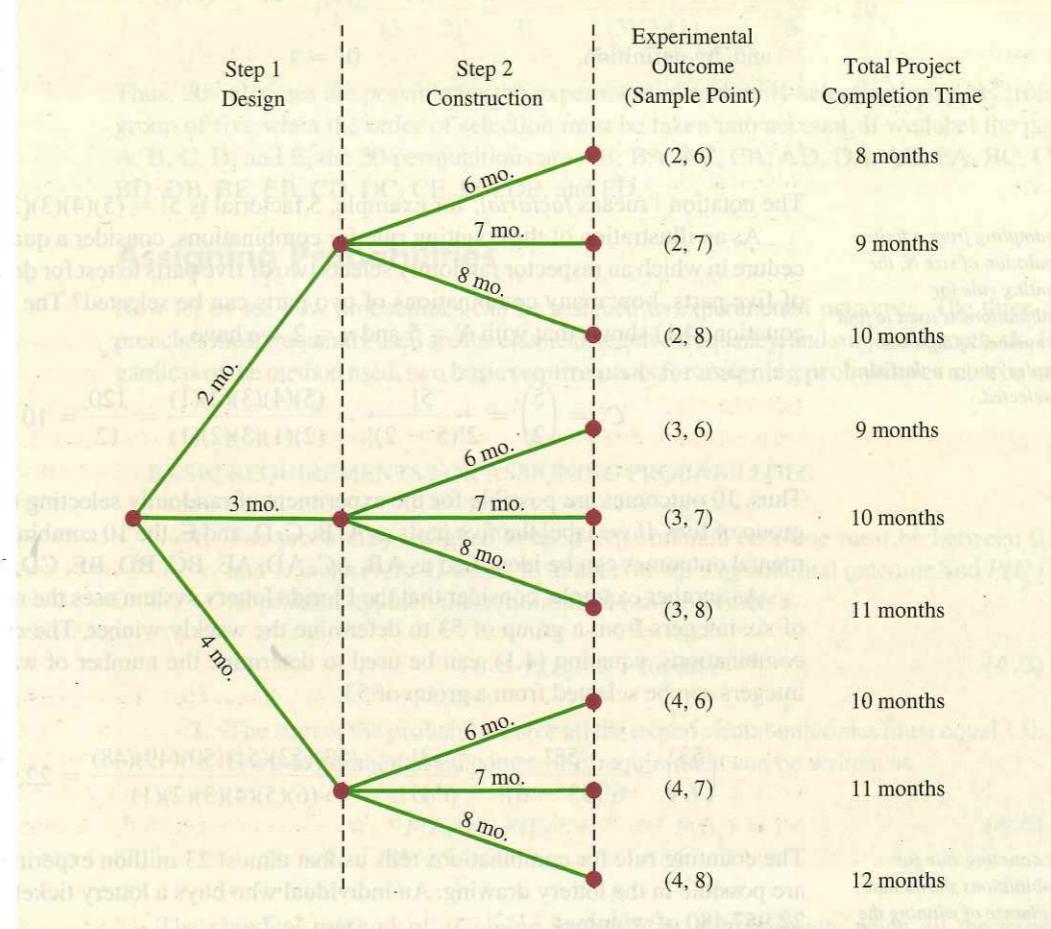
FIGURE 4.3 TREE DIAGRAM FOR THE KP&L PROJECT

Figure 4.3, we see that the project will be completed in 8 to 12 months, with six of the nine experimental outcomes providing the desired completion time of 10 months or less. Even though identifying the experimental outcomes may be helpful, we need to consider how probability values can be assigned to the experimental outcomes before making an assessment of the probability that the project will be completed within the desired 10 months.

Combinations A second useful counting rule allows one to count the number of experimental outcomes when the experiment involves selecting n objects from a (usually larger) set of N objects. It is called the counting rule for combinations.

COUNTING RULE FOR COMBINATIONS

The number of combinations of N objects taken n at a time is

$$C_n^N = \binom{N}{n} = \frac{N!}{n!(N-n)!} \quad (4.1)$$

where

$$N! = N(N-1)(N-2)\cdots(2)(1)$$

$$n! = n(n-1)(n-2)\cdots(2)(1)$$

and, by definition,

$$0! = 1$$

The notation ! means *factorial*; for example, 5 factorial is $5! = (5)(4)(3)(2)(1) = 120$.

As an illustration of the counting rule for combinations, consider a quality control procedure in which an inspector randomly selects two of five parts to test for defects. In a group of five parts, how many combinations of two parts can be selected? The counting rule in equation (4.1) shows that with $N = 5$ and $n = 2$, we have

$$C_2^5 = \binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{(5)(4)(3)(2)(1)}{(2)(1)(3)(2)(1)} = \frac{120}{12} = 10$$

Thus, 10 outcomes are possible for the experiment of randomly selecting two parts from a group of five. If we label the five parts as A, B, C, D, and E, the 10 combinations or experimental outcomes can be identified as AB, AC, AD, AE, BC, BD, BE, CD, CE, and DE.

As another example, consider that the Florida lottery system uses the random selection of six integers from a group of 53 to determine the weekly winner. The counting rule for combinations, equation (4.1), can be used to determine the number of ways six different integers can be selected from a group of 53.

$$\binom{53}{6} = \frac{53!}{6!(53-6)!} = \frac{53!}{6!47!} = \frac{(53)(52)(51)(50)(49)(48)}{(6)(5)(4)(3)(2)(1)} = 22,957,480$$

In sampling from a finite population of size N , the counting rule for combinations is used to find the number of different samples of size n that can be selected.

The counting rule for combinations shows that the chance of winning the lottery is very unlikely.

The counting rule for combinations tells us that almost 23 million experimental outcomes are possible in the lottery drawing. An individual who buys a lottery ticket has 1 chance in 22,957,480 of winning.

Permutations A third counting rule that is sometimes useful is the counting rule for permutations. It allows one to compute the number of experimental outcomes when n objects are to be selected from a set of N objects where the order of selection is

important. The same n objects selected in a different order are considered a different experimental outcome.

COUNTING RULE FOR PERMUTATIONS

The number of permutations of N objects taken n at a time is given by

$$P_n^N = n! \binom{N}{n} = \frac{N!}{(N-n)!} \quad (4.2)$$

The counting rule for permutations closely relates to the one for combinations; however, an experiment results in more permutations than combinations for the same number of objects because every selection of n objects can be ordered in $n!$ different ways.

As an example, consider again the quality control process in which an inspector selects two of five parts to inspect for defects. How many permutations may be selected? The counting rule in equation (4.2) shows that with $N = 5$ and $n = 2$, we have

$$P_2^5 = \frac{5!}{(5-2)!} = \frac{5!}{3!} = \frac{(5)(4)(3)(2)(1)}{(3)(2)(1)} = \frac{120}{6} = 20$$

Thus, 20 outcomes are possible for the experiment of randomly selecting two parts from a group of five when the order of selection must be taken into account. If we label the parts A, B, C, D, and E, the 20 permutations are AB, BA, AC, CA, AD, DA, AE, EA, BC, CB, BD, DB, BE, EB, CD, DC, CE, EC, DE, and ED.

Assigning Probabilities

Now let us see how probabilities can be assigned to experimental outcomes. The three approaches most frequently used are the classical, relative frequency, and subjective methods. Regardless of the method used, two **basic requirements for assigning probabilities** must be met.

BASIC REQUIREMENTS FOR ASSIGNING PROBABILITIES

1. The probability assigned to each experimental outcome must be between 0 and 1, inclusively. If we let E_i denote the i th experimental outcome and $P(E_i)$ its probability, then this requirement can be written as

$$0 \leq P(E_i) \leq 1 \text{ for all } i \quad (4.3)$$

2. The sum of the probabilities for all the experimental outcomes must equal 1.0. For n experimental outcomes, this requirement can be written as

$$P(E_1) + P(E_2) + \cdots + P(E_n) = 1 \quad (4.4)$$

The **classical method** of assigning probabilities is appropriate when all the experimental outcomes are equally likely. If n experimental outcomes are possible, a probability of $1/n$ is assigned to each experimental outcome. When using this approach, the two basic requirements for assigning probabilities are automatically satisfied.

For an example, consider the experiment of tossing a fair coin; the two experimental outcomes—head and tail—are equally likely. Because one of the two equally likely outcomes is a head, the probability of observing a head is $1/2$, or .50. Similarly, the probability of observing a tail is also $1/2$, or .50.

As another example, consider the experiment of rolling a die. It would seem reasonable to conclude that the six possible outcomes are equally likely, and hence each outcome is assigned a probability of $1/6$. If $P(1)$ denotes the probability that one dot appears on the upward face of the die, then $P(1) = 1/6$. Similarly, $P(2) = 1/6$, $P(3) = 1/6$, $P(4) = 1/6$, $P(5) = 1/6$, and $P(6) = 1/6$. Note that these probabilities satisfy the two basic requirements of equations (4.3) and (4.4) because each of the probabilities is greater than or equal to zero and they sum to 1.0.

The **relative frequency method** of assigning probabilities is appropriate when data are available to estimate the proportion of the time the experimental outcome will occur if the experiment is repeated a large number of times. As an example, consider a study of waiting times in the X-ray department for a local hospital. A clerk recorded the number of patients waiting for service at 9:00 A.M. on 20 successive days and obtained the following results.

Number Waiting	Number of Days Outcome Occurred
0	2
1	5
2	6
3	4
4	3
Total	20

These data show that on 2 of the 20 days, zero patients were waiting for service; on 5 of the days, one patient was waiting for service; and so on. Using the relative frequency method, we would assign a probability of $2/20 = .10$ to the experimental outcome of zero patients waiting for service, $5/20 = .25$ to the experimental outcome of one patient waiting, $6/20 = .30$ to two patients waiting, $4/20 = .20$ to three patients waiting, and $3/20 = .15$ to four patients waiting. As with the classical method, using the relative frequency method automatically satisfies the two basic requirements of equations (4.3) and (4.4).

The **subjective method** of assigning probabilities is most appropriate when one cannot realistically assume that the experimental outcomes are equally likely and when little relevant data are available. When the subjective method is used to assign probabilities to the experimental outcomes, we may use any information available, such as our experience or intuition. After considering all available information, a probability value that expresses our *degree of belief* (on a scale from 0 to 1) that the experimental outcome will occur is specified. Because subjective probability expresses a person's degree of belief, it is personal. Using the subjective method, different people can be expected to assign different probabilities to the same experimental outcome.

The subjective method requires extra care to ensure that the two basic requirements of equations (4.3) and (4.4) are satisfied. Regardless of a person's degree of belief, the probability value assigned to each experimental outcome must be between 0 and 1, inclusive, and the sum of all the probabilities for the experimental outcomes must equal 1.0.

Consider the case in which Tom and Judy Elsbernd make an offer to purchase a house. Two outcomes are possible:

$$\begin{aligned} E_1 &= \text{their offer is accepted} \\ E_2 &= \text{their offer is rejected} \end{aligned}$$

Judy believes that the probability their offer will be accepted is .8; thus, Judy would set $P(E_1) = .8$ and $P(E_2) = .2$. Tom, however, believes that the probability that their offer will be accepted is .6; hence, Tom would set $P(E_1) = .6$ and $P(E_2) = .4$. Note that Tom's probability estimate for E_1 reflects a greater pessimism that their offer will be accepted.

Both Judy and Tom assigned probabilities that satisfy the two basic requirements. The fact that their probability estimates are different emphasizes the personal nature of the subjective method.

Even in business situations where either the classical or the relative frequency approach can be applied, managers may want to provide subjective probability estimates. In such cases, the best probability estimates often are obtained by combining the estimates from the classical or relative frequency approach with subjective probability estimates.

Bayes' theorem (see Section 4.5) provides a means for combining subjectively determined prior probabilities with probabilities obtained by other means to obtain revised, or posterior, probabilities.

Probabilities for the KP&L Project

To perform further analysis on the KP&L project, we must develop probabilities for each of the nine experimental outcomes listed in Table 4.1. On the basis of experience and judgment, management concluded that the experimental outcomes were not equally likely. Hence, the classical method of assigning probabilities could not be used. Management then decided to conduct a study of the completion times for similar projects undertaken by KP&L over the past three years. The results of a study of 40 similar projects are summarized in Table 4.2.

After reviewing the results of the study, management decided to employ the relative frequency method of assigning probabilities. Management could have provided subjective probability estimates, but felt that the current project was quite similar to the 40 previous projects. Thus, the relative frequency method was judged best.

In using the data in Table 4.2 to compute probabilities, we note that outcome (2, 6)—stage 1 completed in 2 months and stage 2 completed in 6 months—occurred six times in the 40 projects. We can use the relative frequency method to assign a probability of $6/40 = .15$ to this outcome. Similarly, outcome (2, 7) also occurred in six of the 40 projects, providing a $6/40 = .15$ probability. Continuing in this manner, we obtain the probability assignments for the sample points of the KP&L project shown in Table 4.3. Note that $P(2, 6)$ represents the probability of the sample point (2, 6), $P(2, 7)$ represents the probability of the sample point (2, 7), and so on.

TABLE 4.2 COMPLETION RESULTS FOR 40 KP&L PROJECTS

Completion Time (months)	Number of Past Projects Having These Completion Times		
Stage 1 Design	Stage 2 Construction	Sample Point	
2	6	(2, 6)	6
2	7	(2, 7)	6
2	8	(2, 8)	2
3	6	(3, 6)	4
3	7	(3, 7)	8
3	8	(3, 8)	2
4	6	(4, 6)	2
4	7	(4, 7)	4
4	8	(4, 8)	6
		Total	40

TABLE 4.3 PROBABILITY ASSIGNMENTS FOR THE KP&L PROJECT BASED ON THE RELATIVE FREQUENCY METHOD

Sample Point	Project Completion Time	Probability of Sample Point
(2, 6)	8 months	$P(2, 6) = 6/40 = .15$
(2, 7)	9 months	$P(2, 7) = 6/40 = .15$
(2, 8)	10 months	$P(2, 8) = 2/40 = .05$
(3, 6)	9 months	$P(3, 6) = 4/40 = .10$
(3, 7)	10 months	$P(3, 7) = 8/40 = .20$
(3, 8)	11 months	$P(3, 8) = 2/40 = .05$
(4, 6)	10 months	$P(4, 6) = 2/40 = .05$
(4, 7)	11 months	$P(4, 7) = 4/40 = .10$
(4, 8)	12 months	$P(4, 8) = 6/40 = .15$
	Total	1.00

NOTES AND COMMENTS

- In statistics, the notion of an experiment differs somewhat from the notion of an experiment in the physical sciences. In the physical sciences, researchers usually conduct an experiment in a laboratory or a controlled environment in order to learn about cause and effect. In statistical experiments, probability determines outcomes. Even though the experiment is repeated in exactly the same way, an entirely different

outcome may occur. Because of this influence of probability on the outcome, the experiments of statistics are sometimes called *random experiments*.

- When drawing a random sample without replacement from a population of size N , the counting rule for combinations is used to find the number of different samples of size n that can be selected.

Exercises

Methods

- An experiment has three steps with three outcomes possible for the first step, two outcomes possible for the second step, and four outcomes possible for the third step. How many experimental outcomes exist for the entire experiment?
- How many ways can three items be selected from a group of six items? Use the letters A, B, C, D, E, and F to identify the items, and list each of the different combinations of three items.
- How many permutations of three items can be selected from a group of six? Use the letters A, B, C, D, E, and F to identify the items, and list each of the permutations of items B, D, and F.
- Consider the experiment of tossing a coin three times.
 - Develop a tree diagram for the experiment.
 - List the experimental outcomes.
 - What is the probability for each experimental outcome?
- Suppose an experiment has five equally likely outcomes: E_1, E_2, E_3, E_4, E_5 . Assign probabilities to each outcome and show that the requirements in equations (4.3) and (4.4) are satisfied. What method did you use?
- An experiment with three outcomes has been repeated 50 times, and it was learned that E_1 occurred 20 times, E_2 occurred 13 times, and E_3 occurred 17 times. Assign probabilities to the outcomes. What method did you use?

SELF test

4.1 Experiments, Counting Rules, and Assigning Probabilities

- A decision maker subjectively assigned the following probabilities to the four outcomes of an experiment: $P(E_1) = .10$, $P(E_2) = .15$, $P(E_3) = .40$, and $P(E_4) = .20$. Are these probability assignments valid? Explain.

Applications

- In the city of Milford, applications for zoning changes go through a two-step process: a review by the planning commission and a final decision by the city council. At step 1 the planning commission reviews the zoning change request and makes a positive or negative recommendation concerning the change. At step 2 the city council reviews the planning commission's recommendation and then votes to approve or to disapprove the zoning change. Suppose the developer of an apartment complex submits an application for a zoning change. Consider the application process as an experiment.
 - How many sample points are there for this experiment? List the sample points.
 - Construct a tree diagram for the experiment.
- Simple random sampling uses a sample of size n from a population of size N to obtain data that can be used to make inferences about the characteristics of a population. Suppose that, from a population of 50 bank accounts, we want to take a random sample of four accounts in order to learn about the population. How many different random samples of four accounts are possible?
- Venture capital can provide a big boost in funds available to companies. According to Venture Economics (*Investor's Business Daily*, April 28, 2000), of 2374 venture capital disbursements, 1434 were to companies in California, 390 were to companies in Massachusetts, 217 were to companies in New York, and 112 were to companies in Colorado. Twenty-two percent of the companies receiving funds were in the early stages of development and 55% of the companies were in an expansion stage. Suppose you want to randomly choose one of these companies to learn about how venture capital funds are used.
 - What is the probability the company chosen will be from California?
 - What is the probability the company chosen will not be from one of the four states mentioned?
 - What is the probability the company will not be in the early stages of development?
 - Assume the companies in the early stages of development were evenly distributed across the country. How many Massachusetts companies receiving venture capital funds were in their early stages of development?
 - The total amount of funds invested was \$32.4 billion. Estimate the amount that went to Colorado.
- The National Highway Traffic Safety Administration (NHTSA) conducted a survey to learn about how drivers throughout the United States are using seat belts (Associated Press, August 25, 2003). Sample data consistent with the NHTSA survey are as follows.

Region	Driver Using Seat Belt?	
	Yes	No
Northeast	148	52
Midwest	162	54
South	296	74
West	252	48
Total	858	228

- For the United States, what is the probability that a driver is using a seat belt?
- The seat belt usage probability for a U.S. driver a year earlier was .75. NHTSA chief Dr. Jeffrey Runge had hoped for a .78 probability in 2003. Would he have been pleased with the 2003 survey results?

SELF test

- c. What is the probability of seat belt usage by region of the country? What region has the highest seat belt usage?
 - d. What proportion of the drivers in the sample came from each region of the country? What region had the most drivers selected? What region had the second most drivers selected?
 - e. Assuming the total number of drivers in each region is the same, do you see any reason why the probability estimate in part (a) might be too high? Explain.
12. The Powerball lottery is played twice each week in 28 states, the Virgin Islands, and the District of Columbia. To play Powerball a participant must purchase a ticket and then select five numbers from the digits 1 through 55 and a Powerball number from the digits 1 through 42. To determine the winning numbers for each game, lottery officials draw five white balls out of a drum with 55 white balls, and one red ball out of a drum with 42 red balls. To win the jackpot, a participant's numbers must match the numbers on the five white balls in any order and the number on the red Powerball. Eight coworkers at the ConAgra Foods plant in Lincoln, Nebraska, claimed the record \$365 million jackpot on February 18, 2006, by matching the numbers 15-17-43-44-49 and the Powerball number 29. A variety of other cash prizes are awarded each time the game is played. For instance, a prize of \$200,000 is paid if the participant's five numbers match the numbers on the five white balls (<http://www.powerball.com>, March 19, 2006).
- a. Compute the number of ways the first five numbers can be selected.
 - b. What is the probability of winning a prize of \$200,000 by matching the numbers on the five white balls?
 - c. What is the probability of winning the Powerball jackpot?
13. A company that manufactures toothpaste is studying five different package designs. Assuming that one design is just as likely to be selected by a consumer as any other design, what selection probability would you assign to each of the package designs? In an actual experiment, 100 consumers were asked to pick the design they preferred. The following data were obtained. Do the data confirm the belief that one design is just as likely to be selected as another? Explain.

Design	Number of Times Preferred
1	5
2	15
3	30
4	40
5	10

4.2

Events and Their Probabilities

In the introduction to this chapter we used the term *event* much as it would be used in everyday language. Then, in Section 4.1 we introduced the concept of an experiment and its associated experimental outcomes or sample points. Sample points and events provide the foundation for the study of probability. As a result, we must now introduce the formal definition of an **event** as it relates to sample points. Doing so will provide the basis for determining the probability of an event.

EVENT

An event is a collection of sample points.

For an example, let us return to the KP&L project and assume that the project manager is interested in the event that the entire project can be completed in 10 months or less. Referring to Table 4.3, we see that six sample points—(2, 6), (2, 7), (2, 8), (3, 6), (3, 7), and (4, 6)—provide a project completion time of 10 months or less. Let C denote the event that the project is completed in 10 months or less; we write

$$C = \{(2, 6), (2, 7), (2, 8), (3, 6), (3, 7), (4, 6)\}$$

Event C is said to occur if *any one* of these six sample points appears as the experimental outcome.

Other events that might be of interest to KP&L management include the following.

$$L = \text{The event that the project is completed in } \textit{less} \text{ than 10 months}$$

$$M = \text{The event that the project is completed in } \textit{more} \text{ than 10 months}$$

Using the information in Table 4.3, we see that these events consist of the following sample points.

$$L = \{(2, 6), (2, 7), (3, 6)\}$$

$$M = \{(3, 8), (4, 7), (4, 8)\}$$

A variety of additional events can be defined for the KP&L project, but in each case the event must be identified as a collection of sample points for the experiment.

Given the probabilities of the sample points shown in Table 4.3, we can use the following definition to compute the probability of any event that KP&L management might want to consider.

PROBABILITY OF AN EVENT

The probability of any event is equal to the sum of the probabilities of the sample points in the event.

Using this definition, we calculate the probability of a particular event by adding the probabilities of the sample points (experimental outcomes) that make up the event. We can now compute the probability that the project will take 10 months or less to complete. Because this event is given by $C = \{(2, 6), (2, 7), (2, 8), (3, 6), (3, 7), (4, 6)\}$, the probability of event C , denoted $P(C)$, is given by

$$P(C) = P(2, 6) + P(2, 7) + P(2, 8) + P(3, 6) + P(3, 7) + P(4, 6)$$

Based on the sample point probabilities in Table 4.3, we have

$$P(C) = .15 + .15 + .05 + .10 + .20 + .05 = .70$$

Similarly, because the event that the project is completed in less than 10 months is given by $L = \{(2, 6), (2, 7), (3, 6)\}$, the probability of this event is given by

$$\begin{aligned} P(L) &= P(2, 6) + P(2, 7) + P(3, 6) \\ &= .15 + .15 + .10 = .40 \end{aligned}$$

Finally, for the event that the project is completed in more than 10 months, we have $M = \{(3, 8), (4, 7), (4, 8)\}$ and thus

$$\begin{aligned} P(M) &= P(3, 8) + P(4, 7) + P(4, 8) \\ &= .05 + .10 + .15 = .30 \end{aligned}$$

Using these probability results, we can now tell KP&L management that there is a .70 probability that the project will be completed in 10 months or less, a .40 probability that the project will be completed in less than 10 months, and a .30 probability that the project will be completed in more than 10 months. This procedure of computing event probabilities can be repeated for any event of interest to the KP&L management.

Any time that we can identify all the sample points of an experiment and assign probabilities to each, we can compute the probability of an event using the definition. However, in many experiments the large number of sample points makes the identification of the sample points, as well as the determination of their associated probabilities, extremely cumbersome, if not impossible. In the remaining sections of this chapter, we present some basic probability relationships that can be used to compute the probability of an event without knowledge of all the sample point probabilities.

NOTES AND COMMENTS

1. The sample space, S , is an event. Because it contains all the experimental outcomes, it has a probability of 1; that is, $P(S) = 1$.
2. When the classical method is used to assign probabilities, the assumption is that the experimental outcomes are equally likely. In

such cases, the probability of an event can be computed by counting the number of experimental outcomes in the event and dividing the result by the total number of experimental outcomes.

Exercises

Methods

14. An experiment has four equally likely outcomes: E_1, E_2, E_3 , and E_4 .
 - a. What is the probability that E_2 occurs?
 - b. What is the probability that any two of the outcomes occur (e.g., E_1 or E_3)?
 - c. What is the probability that any three of the outcomes occur (e.g., E_1 or E_2 or E_4)?
15. Consider the experiment of selecting a playing card from a deck of 52 playing cards. Each card corresponds to a sample point with a $1/52$ probability.
 - a. List the sample points in the event an ace is selected.
 - b. List the sample points in the event a club is selected.
 - c. List the sample points in the event a face card (jack, queen, or king) is selected.
 - d. Find the probabilities associated with each of the events in parts (a), (b), and (c).
16. Consider the experiment of rolling a pair of dice. Suppose that we are interested in the sum of the face values showing on the dice.
 - a. How many sample points are possible? (*Hint:* Use the counting rule for multiple-step experiments.)
 - b. List the sample points.
 - c. What is the probability of obtaining a value of 7?
 - d. What is the probability of obtaining a value of 9 or greater?
 - e. Because each roll has six possible even values (2, 4, 6, 8, 10, and 12) and only five possible odd values (3, 5, 7, 9, and 11), the dice should show even values more often than odd values. Do you agree with this statement? Explain.
 - f. What method did you use to assign the probabilities requested?

SELF test

SELF test

4.2 Events and Their Probabilities

Applications

17. Refer to the KP&L sample points and sample point probabilities in Tables 4.2 and 4.3.
 - a. The design stage (stage 1) will run over budget if it takes 4 months to complete. List the sample points in the event the design stage is over budget.
 - b. What is the probability that the design stage is over budget?
 - c. The construction stage (stage 2) will run over budget if it takes 8 months to complete. List the sample points in the event the construction stage is over budget.
 - d. What is the probability that the construction stage is over budget?
 - e. What is the probability that both stages are over budget?
18. To investigate how often we eat at home as a family during the week, Harris Interactive surveyed 496 adults living with children under the age of 18 (*USA Today*, January 3, 2007). The survey results are shown in the following table.

Family Meals	Number of Times Outcome Occurred
0	11
1	11
2	30
3	36
4	36
5	119
6	114
7 or more	139

For a randomly selected family with children under the age of 18, compute the following:

- a. The probability the family eats no meals at home during the week.
- b. The probability the family eats at least four meals at home during the week.
- c. The probability the family eats two or fewer meals at home during the week.
19. The National Sporting Goods Association conducted a survey of persons 7 years of age or older about participation in sports activities (*Statistical Abstract of the United States: 2002*). The total population in this age group was reported at 248.5 million, with 120.9 million male and 127.6 million female. The number of participants for the top five sports activities appears here.

Activity	Participants (millions)	
	Male	Female
Bicycle riding	22.2	21.0
Camping	25.6	24.3
Exercise walking	28.7	57.7
Exercising with equipment	20.4	24.4
Swimming	26.4	34.4

- a. For a randomly selected female, estimate the probability of participation in each of the sports activities.
- b. For a randomly selected male, estimate the probability of participation in each of the sports activities.
- c. For a randomly selected person, what is the probability the person participates in exercise walking?
- d. Suppose you just happen to see an exercise walker going by. What is the probability the walker is a woman? What is the probability the walker is a man?

20. *Fortune* magazine publishes an annual list of the 500 largest companies in the United States. The following data show the five states with the largest number of *Fortune* 500 companies (*The New York Times Almanac*, 2006).

State	Number of Companies
New York	54
California	52
Texas	48
Illinois	33
Ohio	30

Suppose a *Fortune* 500 company is chosen for a follow-up questionnaire. What are the probabilities of the following events?

- a. Let N be the event the company is headquartered in New York. Find $P(N)$.
 - b. Let T be the event the company is headquartered in Texas. Find $P(T)$.
 - c. Let B be the event the company is headquartered in one of these five states. Find $P(B)$.
21. The U.S. population by age is as follows (*The World Almanac*, 2004). The data are in millions of people.

Age	Number
19 and under	80.5
20 to 24	19.0
25 to 34	39.9
35 to 44	45.2
45 to 54	37.7
55 to 64	24.3
65 and over	35.0

Assume that a person will be randomly chosen from this population.

- a. What is the probability the person is 20 to 24 years old?
- b. What is the probability the person is 20 to 34 years old?
- c. What is the probability the person is 45 years or older?

4.3

Some Basic Relationships of Probability

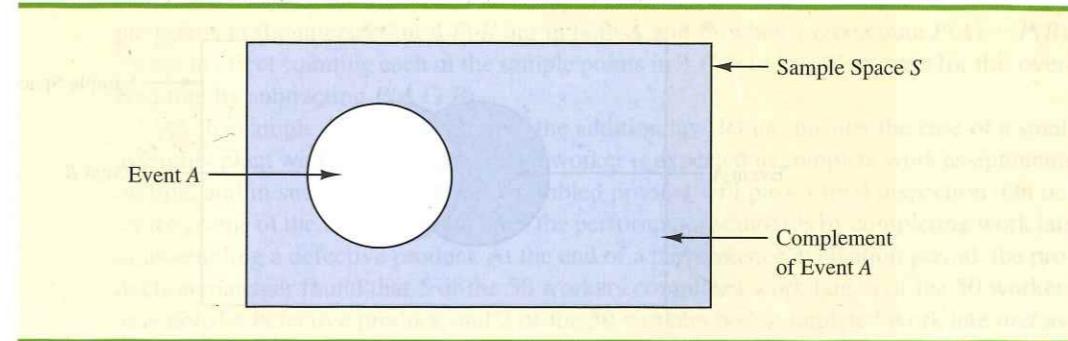
Complement of an Event

Given an event A , the **complement of A** is defined to be the event consisting of all sample points that are *not* in A . The complement of A is denoted by A^c . Figure 4.4 is a diagram, known as a **Venn diagram**, which illustrates the concept of a complement. The rectangular area represents the sample space for the experiment and as such contains all possible sample points. The circle represents event A and contains only the sample points that belong to A . The shaded region of the rectangle contains all sample points not in event A and is by definition the complement of A .

In any probability application, either event A or its complement A^c must occur. Therefore, we have

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

FIGURE 4.4 COMPLEMENT OF EVENT A IS SHADED



Solving for $P(A)$, we obtain the following result.

COMPUTING PROBABILITY USING THE COMPLEMENT

$$P(A) = 1 - P(A^c) \quad (4.5)$$

Equation (4.5) shows that the probability of an event A can be computed easily if the probability of its complement, $P(A^c)$, is known.

As an example, consider the case of a sales manager who, after reviewing sales reports, states that 80% of new customer contacts result in no sale. By allowing A to denote the event of a sale and A^c to denote the event of no sale, the manager is stating that $P(A^c) = .80$. Using equation (4.5), we see that

$$P(A) = 1 - P(A^c) = 1 - .80 = .20$$

We can conclude that a new customer contact has a .20 probability of resulting in a sale.

In another example, a purchasing agent states a .90 probability that a supplier will send a shipment that is free of defective parts. Using the complement, we can conclude that there is a $1 - .90 = .10$ probability that the shipment will contain defective parts.

Addition Law

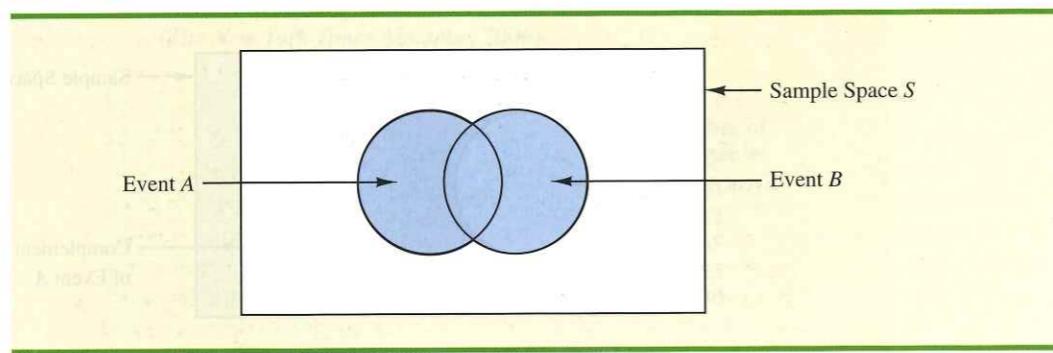
The addition law is helpful when we are interested in knowing the probability that at least one of two events occurs. That is, with events A and B we are interested in knowing the probability that event A or event B or both occur.

Before we present the addition law, we need to discuss two concepts related to the combination of events: the *union* of events and the *intersection* of events. Given two events A and B , the **union of A and B** is defined as follows.

UNION OF TWO EVENTS

The *union* of A and B is the event containing *all* sample points belonging to A or B or both. The union is denoted by $A \cup B$.

The Venn diagram in Figure 4.5 depicts the union of events A and B . Note that the two circles contain all the sample points in event A as well as all the sample points in event B .

FIGURE 4.5 UNION OF EVENTS A AND B IS SHADED

The fact that the circles overlap indicates that some sample points are contained in both A and B .

The definition of the **intersection of A and B** follows.

INTERSECTION OF TWO EVENTS

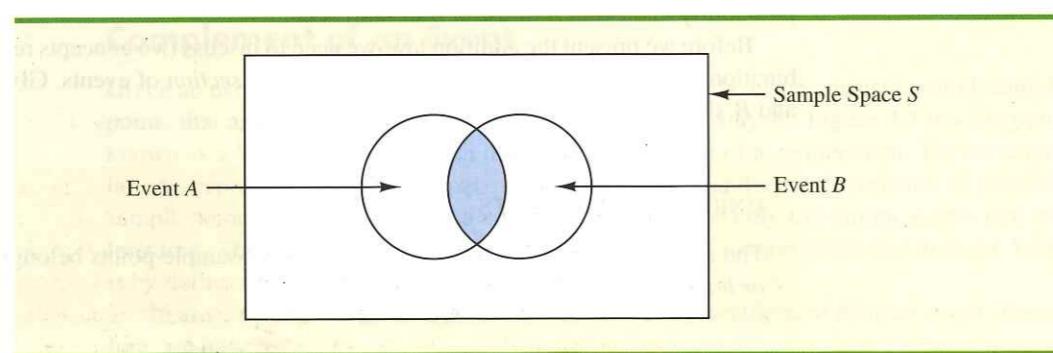
Given two events A and B , the *intersection of A and B* is the event containing the sample points belonging to *both A and B* . The intersection is denoted by $A \cap B$.

The Venn diagram depicting the intersection of events A and B is shown in Figure 4.6. The area where the two circles overlap is the intersection; it contains the sample points that are in both A and B .

Let us now continue with a discussion of the addition law. The **addition law** provides a way to compute the probability that event A or event B or both occur. In other words, the addition law is used to compute the probability of the union of two events. The addition law is written as follows.

ADDITION LAW

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (4.6)$$

FIGURE 4.6 INTERSECTION OF EVENTS A AND B IS SHADED

To understand the addition law intuitively, note that the first two terms in the addition law, $P(A) + P(B)$, account for all the sample points in $A \cup B$. However, because the sample points in the intersection $A \cap B$ are in both A and B , when we compute $P(A) + P(B)$, we are in effect counting each of the sample points in $A \cap B$ twice. We correct for this over-counting by subtracting $P(A \cap B)$.

As an example of an application of the addition law, let us consider the case of a small assembly plant with 50 employees. Each worker is expected to complete work assignments on time and in such a way that the assembled product will pass a final inspection. On occasion, some of the workers fail to meet the performance standards by completing work late or assembling a defective product. At the end of a performance evaluation period, the production manager found that 5 of the 50 workers completed work late, 6 of the 50 workers assembled a defective product, and 2 of the 50 workers both completed work late *and* assembled a defective product.

Let

L = the event that the work is completed late

D = the event that the assembled product is defective

The relative frequency information leads to the following probabilities.

$$P(L) = \frac{5}{50} = .10$$

$$P(D) = \frac{6}{50} = .12$$

$$P(L \cap D) = \frac{2}{50} = .04$$

After reviewing the performance data, the production manager decided to assign a poor performance rating to any employee whose work was either late or defective; thus the event of interest is $L \cup D$. What is the probability that the production manager assigned an employee a poor performance rating?

Note that the probability question is about the union of two events. Specifically, we want to know $P(L \cup D)$. Using equation (4.6), we have

$$P(L \cup D) = P(L) + P(D) - P(L \cap D)$$

Knowing values for the three probabilities on the right side of this expression, we can write

$$P(L \cup D) = .10 + .12 - .04 = .18$$

This calculation tells us that there is a .18 probability that a randomly selected employee received a poor performance rating.

As another example of the addition law, consider a recent study conducted by the personnel manager of a major computer software company. The study showed that 30% of the employees who left the firm within two years did so primarily because they were dissatisfied with their salary, 20% left because they were dissatisfied with their work assignments, and 12% of the former employees indicated dissatisfaction with *both* their salary and their work assignments. What is the probability that an employee who leaves within

two years does so because of dissatisfaction with salary, dissatisfaction with the work assignment, or both?

Let

S = the event that the employee leaves because of salary

W = the event that the employee leaves because of work assignment

We have $P(S) = .30$, $P(W) = .20$, and $P(S \cap W) = .12$. Using equation (4.6), the addition law, we have

$$P(S \cup W) = P(S) + P(W) - P(S \cap W) = .30 + .20 - .12 = .38$$

We find a .38 probability that an employee leaves for salary or work assignment reasons.

Before we conclude our discussion of the addition law, let us consider a special case that arises for **mutually exclusive events**.

MUTUALLY EXCLUSIVE EVENTS

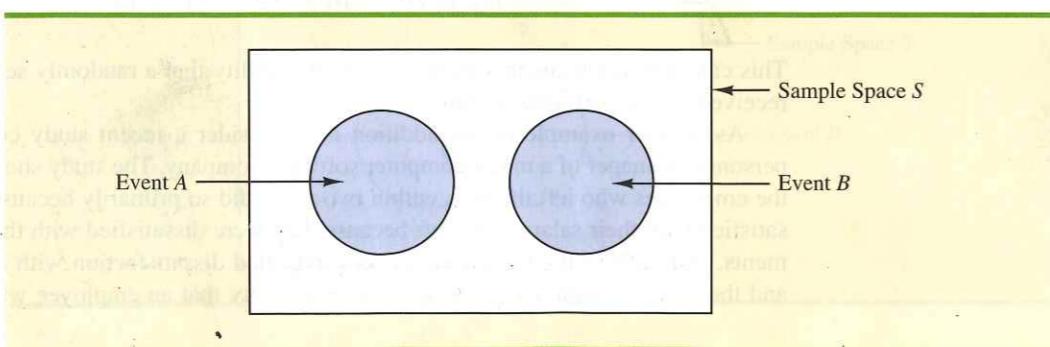
Two events are said to be mutually exclusive if the events have no sample points in common.

Events A and B are mutually exclusive if, when one event occurs, the other cannot occur. Thus, a requirement for A and B to be mutually exclusive is that their intersection must contain no sample points. The Venn diagram depicting two mutually exclusive events A and B is shown in Figure 4.7. In this case $P(A \cap B) = 0$ and the addition law can be written as follows.

ADDITION LAW FOR MUTUALLY EXCLUSIVE EVENTS

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

FIGURE 4.7 MUTUALLY EXCLUSIVE EVENTS



Exercises

Methods

22. Suppose that we have a sample space with five equally likely experimental outcomes: E_1, E_2, E_3, E_4, E_5 . Let

$$A = \{E_1, E_2\}$$

$$B = \{E_3, E_4\}$$

$$C = \{E_2, E_3, E_5\}$$

- a. Find $P(A)$, $P(B)$, and $P(C)$.
 - b. Find $P(A \cup B)$. Are A and B mutually exclusive?
 - c. Find A^c , C^c , $P(A^c)$, and $P(C^c)$.
 - d. Find $A \cup B^c$ and $P(A \cup B^c)$.
 - e. Find $P(B \cup C)$.
23. Suppose that we have a sample space $S = \{E_1, E_2, E_3, E_4, E_5, E_6, E_7\}$, where E_1, E_2, \dots, E_7 denote the sample points. The following probability assignments apply: $P(E_1) = .05$, $P(E_2) = .20$, $P(E_3) = .20$, $P(E_4) = .25$, $P(E_5) = .15$, $P(E_6) = .10$, and $P(E_7) = .05$. Let

$$A = \{E_1, E_4, E_6\}$$

$$B = \{E_2, E_4, E_7\}$$

$$C = \{E_2, E_3, E_5, E_7\}$$

- a. Find $P(A)$, $P(B)$, and $P(C)$.
- b. Find $A \cup B$ and $P(A \cup B)$.
- c. Find $A \cap B$ and $P(A \cap B)$.
- d. Are events A and C mutually exclusive?
- e. Find B^c and $P(B^c)$.

Applications

24. Clarkson University surveyed alumni to learn more about what they think of Clarkson. One part of the survey asked respondents to indicate whether their overall experience at Clarkson fell short of expectations, met expectations, or surpassed expectations. The results showed that 4% of the respondents did not provide a response, 26% said that their experience fell short of expectations, and 65% of the respondents said that their experience met expectations (*Clarkson Magazine*, Summer 2001).

- a. If we chose an alumnus at random, what is the probability that the alumnus would say their experience surpassed expectations?
- b. If we chose an alumnus at random, what is the probability that the alumnus would say their experience met or surpassed expectations?

25. The U.S. Census Bureau provides data on the number of young adults, ages 18–24, who are living in their parents' home.* Let

M = the event a male young adult is living in his parents' home

F = the event a female young adult is living in her parents' home

If we randomly select a male young adult and a female young adult, the Census Bureau data enable us to conclude $P(M) = .56$ and $P(F) = .42$ (*The World Almanac*, 2006). The probability that both are living in their parents' home is .24.

- a. What is the probability at least one of the two young adults selected is living in his or her parents' home?
- b. What is the probability both young adults selected are living on their own (neither is living in their parents' home)?

*The data include single young adults who are living in college dormitories because it is assumed these young adults will return to their parents' home when school is not in session.

SELF test

26. Data on the 30 largest stock and balanced funds provided one-year and five-year percentage returns for the period ending March 31, 2000 (*The Wall Street Journal*, April 10, 2000). Suppose we consider a one-year return in excess of 50% to be high and a five-year return in excess of 300% to be high. Nine of the funds had one-year returns in excess of 50%, seven of the funds had five-year returns in excess of 300%, and five of the funds had both one-year returns in excess of 50% and five-year returns in excess of 300%.
- What is the probability of a high one-year return, and what is the probability of a high five-year return?
 - What is the probability of both a high one-year return and a high five-year return?
 - What is the probability of neither a high one-year return nor a high five-year return?
27. A 2001 preseason NCAA football poll asked respondents to answer the question, "Will the Big Ten or the Pac-10 have a team in this year's national championship game, the Rose Bowl?" Of the 13,429 respondents, 2961 said the Big Ten would, 4494 said the Pac-10 would, and 6823 said neither the Big Ten nor the Pac-10 would have a team in the Rose Bowl (<http://www.yahoo.com>, August 30, 2001).
- What is the probability that a respondent said neither the Big Ten nor the Pac-10 would have a team in the Rose Bowl?
 - What is the probability that a respondent said either the Big Ten or the Pac-10 would have a team in the Rose Bowl?
 - Find the probability that a respondent said both the Big Ten and the Pac-10 would have a team in the Rose Bowl.
28. A survey of magazine subscribers showed that 45.8% rented a car during the past 12 months for business reasons, 54% rented a car during the past 12 months for personal reasons, and 30% rented a car during the past 12 months for both business and personal reasons.
- What is the probability that a subscriber rented a car during the past 12 months for business or personal reasons?
 - What is the probability that a subscriber did not rent a car during the past 12 months for either business or personal reasons?
29. High school seniors with strong academic records apply to the nation's most selective colleges in greater numbers each year. Because the number of slots remains relatively stable, some colleges reject more early applicants. The University of Pennsylvania received 2851 applications for early admission. Of this group, it admitted 1033 students early, rejected 854 outright, and deferred 964 to the regular admissions pool for further consideration. In the past, Penn has admitted about 18% of the deferred early admission applicants during the regular admission process. Counting the additional students who were admitted during the regular admission process, the total class size was 2375 (*USA Today*, January 24, 2001). Let E , R , and D represent the events that a student who applies for early admission is admitted early, rejected outright, or deferred to the regular admissions pool.
- Use the data to estimate $P(E)$, $P(R)$, and $P(D)$.
 - Are events E and D mutually exclusive? Find $P(E \cap D)$.
 - For the 2375 students admitted to Penn, what is the probability that a randomly selected student was accepted for early admission?
 - Suppose a student applies to Penn for early admission. What is the probability the student will be admitted for early admission or be deferred and later admitted during the regular admission process?

4.4**Conditional Probability**

Often, the probability of an event is influenced by whether a related event already occurred. Suppose we have an event A with probability $P(A)$. If we obtain new information and learn that a related event, denoted by B , already occurred, we will want to take advantage of this

information by calculating a new probability for event A . This new probability of event A is called a **conditional probability** and is written $P(A | B)$. We use the notation $|$ to indicate that we are considering the probability of event A given the condition that event B has occurred. Hence, the notation $P(A | B)$ reads "the probability of A given B ."

As an illustration of the application of conditional probability, consider the situation of the promotion status of male and female officers of a major metropolitan police force in the eastern United States. The police force consists of 1200 officers, 960 men and 240 women. Over the past two years, 324 officers on the police force received promotions. The specific breakdown of promotions for male and female officers is shown in Table 4.4.

After reviewing the promotion record, a committee of female officers raised a discrimination case on the basis that 288 male officers had received promotions but only 36 female officers had received promotions. The police administration argued that the relatively low number of promotions for female officers was due not to discrimination, but to the fact that relatively few females are members of the police force. Let us show how conditional probability could be used to analyze the discrimination charge.

Let

$$\begin{aligned} M &= \text{event an officer is a man} \\ W &= \text{event an officer is a woman} \\ A &= \text{event an officer is promoted} \\ A^c &= \text{event an officer is not promoted} \end{aligned}$$

Dividing the data values in Table 4.4 by the total of 1200 officers enables us to summarize the available information with the following probability values.

$$P(M \cap A) = 288/1200 = .24 = \text{probability that a randomly selected officer is a man and is promoted}$$

$$P(M \cap A^c) = 672/1200 = .56 = \text{probability that a randomly selected officer is a man and is not promoted}$$

$$P(W \cap A) = 36/1200 = .03 = \text{probability that a randomly selected officer is a woman and is promoted}$$

$$P(W \cap A^c) = 204/1200 = .17 = \text{probability that a randomly selected officer is a woman and is not promoted}$$

Because each of these values gives the probability of the intersection of two events, the probabilities are called **joint probabilities**. Table 4.5, which provides a summary of the probability information for the police officer promotion situation, is referred to as a *joint probability table*.

The values in the margins of the joint probability table provide the probabilities of each event separately. That is, $P(M) = .80$, $P(W) = .20$, $P(A) = .27$, and $P(A^c) = .73$. These

TABLE 4.4 PROMOTION STATUS OF POLICE OFFICERS OVER THE PAST TWO YEARS

	Men	Women	Total
Promoted	288	36	324
Not Promoted	672	204	876
Total	960	240	1200

TABLE 4.5 JOINT PROBABILITY TABLE FOR PROMOTIONS

	Men (M)	Women (W)	Total
Promoted (A)	.24	.03	.27
Not Promoted (A^c)	.56	.17	.73
Total	.80	.20	1.00

Marginal probabilities appear in the margins of the table.

probabilities are referred to as **marginal probabilities** because of their location in the margins of the joint probability table. We note that the marginal probabilities are found by summing the joint probabilities in the corresponding row or column of the joint probability table. For instance, the marginal probability of being promoted is $P(A) = P(M \cap A) + P(W \cap A) = .24 + .03 = .27$. From the marginal probabilities, we see that 80% of the force is male, 20% of the force is female, 27% of all officers received promotions, and 73% were not promoted.

Let us begin the conditional probability analysis by computing the probability that an officer is promoted given that the officer is a man. In conditional probability notation, we are attempting to determine $P(A | M)$. To calculate $P(A | M)$, we first realize that this notation simply means that we are considering the probability of the event A (promotion) given that the condition designated as event M (the officer is a man) is known to exist. Thus $P(A | M)$ tells us that we are now concerned only with the promotion status of the 960 male officers. Because 288 of the 960 male officers received promotions, the probability of being promoted given that the officer is a man is $288/960 = .30$. In other words, given that an officer is a man, that officer had a 30% chance of receiving a promotion over the past two years.

This procedure was easy to apply because the values in Table 4.4 show the number of officers in each category. We now want to demonstrate how conditional probabilities such as $P(A | M)$ can be computed directly from related event probabilities rather than the frequency data of Table 4.4.

We have shown that $P(A | M) = 288/960 = .30$. Let us now divide both the numerator and denominator of this fraction by 1200, the total number of officers in the study.

$$P(A | M) = \frac{288}{960} = \frac{288/1200}{960/1200} = \frac{.24}{.80} = .30$$

We now see that the conditional probability $P(A | M)$ can be computed as $.24/.80$. Refer to the joint probability table (Table 4.5). Note in particular that .24 is the joint probability of A and M ; that is, $P(A \cap M) = .24$. Also note that .80 is the marginal probability that a randomly selected officer is a man; that is, $P(M) = .80$. Thus, the conditional probability $P(A | M)$ can be computed as the ratio of the joint probability $P(A \cap M)$ to the marginal probability $P(M)$.

$$P(A | M) = \frac{P(A \cap M)}{P(M)} = \frac{.24}{.80} = .30$$

4.4 Conditional Probability

The fact that conditional probabilities can be computed as the ratio of a joint probability to a marginal probability provides the following general formula for conditional probability calculations for two events A and B .

CONDITIONAL PROBABILITY

$$P(A | B) = \frac{P(A \cap B)}{P(B)}$$
(4.7)

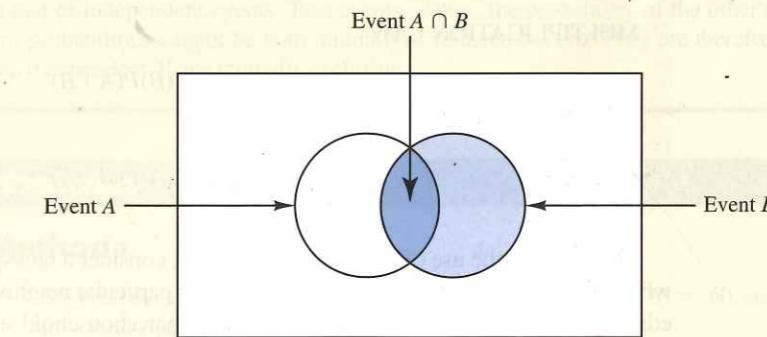
or

$$P(B | A) = \frac{P(A \cap B)}{P(A)}$$
(4.8)

The Venn diagram in Figure 4.8 helps to explain conditional probability. The circle on the right shows that event B has occurred; the portion of the circle that overlaps with event A denotes the event $(A \cap B)$. We know that once event B has occurred, the only way that we can also observe event A is for the event $(A \cap B)$ to occur. Thus, the ratio $P(A \cap B)/P(B)$ provides the conditional probability that we will observe event A given that event B has already occurred.

Let us return to the issue of discrimination against the female officers. The marginal probability in row 1 of Table 4.5 shows that the probability of promotion of an officer is $P(A) = .27$ (regardless of whether that officer is male or female). However, the critical issue in the discrimination case involves the two conditional probabilities $P(A | M)$ and $P(A | W)$. That is, what is the probability of a promotion *given* that the officer is a man, and what is the probability of a promotion *given* that the officer is a woman? If these two probabilities are equal, a discrimination argument has no basis because the chances of a promotion are the same for male and female officers. However, a difference in the two conditional probabilities will support the position that male and female officers are treated differently in promotion decisions.

We already determined that $P(A | M) = .30$. Let us now use the probability values in Table 4.5 and the basic relationship of conditional probability in equation (4.7) to compute

FIGURE 4.8 CONDITIONAL PROBABILITY $P(A | B) = P(A \cap B)/P(B)$ 

the probability that an officer is promoted given that the officer is a woman; that is, $P(A | W)$. Using equation (4.7), with W replacing B , we obtain

$$P(A | W) = \frac{P(A \cap W)}{P(W)} = \frac{.03}{.20} = .15$$

What conclusion do you draw? The probability of a promotion given that the officer is a man is .30, twice the .15 probability of a promotion given that the officer is a woman. Although the use of conditional probability does not in itself prove that discrimination exists in this case, the conditional probability values support the argument presented by the female officers.

Independent Events

In the preceding illustration, $P(A) = .27$, $P(A | M) = .30$, and $P(A | W) = .15$. We see that the probability of a promotion (event A) is affected or influenced by whether the officer is a man or a woman. Particularly, because $P(A | M) \neq P(A)$, we would say that events A and M are dependent events. That is, the probability of event A (promotion) is altered or affected by knowing that event M (the officer is a man) exists. Similarly, with $P(A | W) \neq P(A)$, we would say that events A and W are *dependent events*. However, if the probability of event A is not changed by the existence of event M —that is, $P(A | M) = P(A)$ —we would say that events A and M are **independent events**. This situation leads to the following definition of the independence of two events.

INDEPENDENT EVENTS

Two events A and B are independent if

$$P(A \cap B) = P(A) \quad (4.9)$$

or

$$P(B | A) = P(B) \quad (4.10)$$

Otherwise, the events are dependent.

Multiplication Law

Whereas the addition law of probability is used to compute the probability of a union of two events, the multiplication law is used to compute the probability of the intersection of two events. The multiplication law is based on the definition of conditional probability. Using equations (4.7) and (4.8) and solving for $P(A \cap B)$, we obtain the **multiplication law**.

MULTIPLICATION LAW

$$P(A \cap B) = P(B)P(A | B) \quad (4.11)$$

or

$$P(A \cap B) = P(A)P(B | A) \quad (4.12)$$

To illustrate the use of the multiplication law, consider a newspaper circulation department where it is known that 84% of the households in a particular neighborhood subscribe to the daily edition of the paper. If we let D denote the event that a household subscribes to the daily edition, $P(D) = .84$. In addition, it is known that the probability that a household that already holds a daily subscription also subscribes to the Sunday edition (event S) is .75; that is, $P(S | D) = .75$.

What is the probability that a household subscribes to both the Sunday and daily editions of the newspaper? Using the multiplication law, we compute the desired $P(S \cap D)$ as

$$P(S \cap D) = P(D)P(S | D) = .84(.75) = .63$$

We now know that 63% of the households subscribe to both the Sunday and daily editions.

Before concluding this section, let us consider the special case of the multiplication law when the events involved are independent. Recall that events A and B are independent whenever $P(A | B) = P(A)$ or $P(B | A) = P(B)$. Hence, using equations (4.11) and (4.12) for the special case of independent events, we obtain the following multiplication law.

MULTIPLICATION LAW FOR INDEPENDENT EVENTS

$$P(A \cap B) = P(A)P(B) \quad (4.13)$$

To compute the probability of the intersection of two independent events, we simply multiply the corresponding probabilities. Note that the multiplication law for independent events provides another way to determine whether A and B are independent. That is, if $P(A \cap B) = P(A)P(B)$, then A and B are independent; if $P(A \cap B) \neq P(A)P(B)$, then A and B are dependent.

As an application of the multiplication law for independent events, consider the situation of a service station manager who knows from past experience that 80% of the customers use a credit card when they purchase gasoline. What is the probability that the next two customers purchasing gasoline will each use a credit card? If we let

A = the event that the first customer uses a credit card

B = the event that the second customer uses a credit card

then the event of interest is $A \cap B$. Given no other information, we can reasonably assume that A and B are independent events. Thus,

$$P(A \cap B) = P(A)P(B) = (.80)(.80) = .64$$

To summarize this section, we note that our interest in conditional probability is motivated by the fact that events are often related. In such cases, we say the events are dependent and the conditional probability formulas in equations (4.7) and (4.8) must be used to compute the event probabilities. If two events are not related, they are independent; in this case neither event's probability is affected by whether the other event occurred.

NOTES AND COMMENTS

Do not confuse the notion of mutually exclusive events with that of independent events. Two events with nonzero probabilities cannot be both mutually exclusive and independent. If one mutually exclusive

event is known to occur, the other cannot occur; thus, the probability of the other event occurring is reduced to zero. They are therefore dependent.

Exercises

Methods

30. Suppose that we have two events, A and B , with $P(A) = .50$, $P(B) = .60$, and $P(A \cap B) = .40$.
- Find $P(A | B)$.
 - Find $P(B | A)$.
 - Are A and B independent? Why or why not?



31. Assume that we have two events, A and B , that are mutually exclusive. Assume further that we know $P(A) = .30$ and $P(B) = .40$.
- What is $P(A \cap B)$?
 - What is $P(A | B)$?
 - A student in statistics argues that the concepts of mutually exclusive events and independent events are really the same, and that if events are mutually exclusive they must be independent. Do you agree with this statement? Use the probability information in this problem to justify your answer.
 - What general conclusion would you make about mutually exclusive and independent events given the results of this problem?

Applications

32. Due to rising health insurance costs, 43 million people in the United States go without health insurance (*Time*, December 1, 2003). Sample data representative of the national health insurance coverage are shown here.

		Health Insurance	
		Yes	No
Age	18 to 34	750	170
	35 and older	950	130

- Develop a joint probability table for these data and use the table to answer the remaining questions.
 - What do the marginal probabilities tell you about the age of the U.S. population?
 - What is the probability that a randomly selected individual does not have health insurance coverage?
 - If the individual is between the ages of 18 and 34, what is the probability that the individual does not have health insurance coverage?
 - If the individual is age 35 or older, what is the probability that the individual does not have health insurance coverage?
 - If the individual does not have health insurance, what is the probability that the individual is in the 18 to 34 age group?
 - What does the probability information tell you about health insurance coverage in the United States?
33. In a survey of MBA students, the following data were obtained on "students' first reason for application to the school in which they matriculated."

		Reason for Application			Totals
Enrollment Status	School Quality	School Cost or Convenience	Other		
Full Time	421	393	76	890	
Part Time	400	593	46	1039	
Totals	821	986	122	1929	

- Develop a joint probability table for these data.
- Use the marginal probabilities of school quality, school cost or convenience, and other to comment on the most important reason for choosing a school.

SELF test

- If a student goes full time, what is the probability that school quality is the first reason for choosing a school?
 - If a student goes part time, what is the probability that school quality is the first reason for choosing a school?
 - Let A denote the event that a student is full time and let B denote the event that the student lists school quality as the first reason for applying. Are events A and B independent? Justify your answer.
34. The U.S. Department of Transportation reported that during November, 83.4% of Southwest Airlines flights, 75.1% of US Airways flights, and 70.1% of JetBlue flights arrived on time (*USA Today*, January 4, 2007). Assume that this on-time performance is applicable for flights arriving at concourse A of the Rochester International Airport, and that 40% of the arrivals at concourse A are Southwest Airlines flights, 35% are US Airways flights, and 25% are JetBlue flights.
- Develop a joint probability table with three rows (airlines) and two columns (on-time arrivals vs. late arrivals).
 - An announcement has just been made that Flight 1424 will be arriving at gate 20 in concourse A. What is the most likely airline for this arrival?
 - What is the probability that Flight 1424 will arrive on time?
 - Suppose that an announcement is made saying that Flight 1424 will be arriving late. What is the most likely airline for this arrival? What is the least likely airline?
35. The U.S. Bureau of Labor Statistics collected data on the occupations of workers 25 to 64 years old. The following table shows the number of male and female workers (in millions) in each occupation category (*Statistical Abstract of the United States: 2002*).

Occupation	Male	Female
Managerial/Professional	19079	19021
Tech./Sales/Administrative	11079	19315
Service	4977	7947
Precision Production	11682	1138
Operators/Fabricators/Labor	10576	3482
Farming/Forestry/Fishing	1838	514

- Develop a joint probability table.
 - What is the probability of a female worker being a manager or professional?
 - What is the probability of a male worker being in precision production?
 - Is occupation independent of gender? Justify your answer with a probability calculation.
36. Reggie Miller of the Indiana Pacers is the National Basketball Association's best career free throw shooter, making 89% of his shots (*USA Today*, January 22, 2004). Assume that late in a basketball game, Reggie Miller is fouled and is awarded two shots.
- What is the probability that he will make both shots?
 - What is the probability that he will make at least one shot?
 - What is the probability that he will miss both shots?
 - Late in a basketball game, a team often intentionally fouls an opposing player in order to stop the game clock. The usual strategy is to intentionally foul the other team's worst free throw shooter. Assume that the Indiana Pacers' center makes 58% of his free throw shots. Calculate the probabilities for the center as shown in parts (a), (b), and (c), and show that intentionally fouling the Indiana Pacers' center is a better strategy than intentionally fouling Reggie Miller.
37. Visa Card USA studied how frequently young consumers, ages 18 to 24, use plastic (debit and credit) cards in making purchases (Associated Press, January 16, 2006). The results of the study provided the following probabilities.

- The probability that a consumer uses a plastic card when making a purchase is .37.
- Given that the consumer uses a plastic card, there is a .19 probability that the consumer is 18 to 24 years old.
- Given that the consumer uses a plastic card, there is a .81 probability that the consumer is more than 24 years old.

- U.S. Census Bureau data show that 14% of the consumer population is 18 to 24 years old.
- Given the consumer is 18 to 24 years old, what is the probability that the consumer uses a plastic card?
 - Given the consumer is over 24 years old, what is the probability that the consumer uses a plastic card?
 - What is the interpretation of the probabilities shown in parts (a) and (b)?
 - Should companies such as Visa, MasterCard, and Discover make plastic cards available to the 18 to 24 years old age group before these consumers have had time to establish a credit history? If no, why? If yes, what restrictions might the companies place on this age group?
38. A Morgan Stanley Consumer Research Survey sampled men and women and asked each whether they preferred to drink plain bottled water or a sports drink such as Gatorade or Propel Fitness water (*The Atlanta Journal-Constitution*, December 28, 2005). Suppose 200 men and 200 women participated in the study, and 280 reported they preferred plain bottled water. Of the group preferring a sports drink, 80 were men and 40 were women.

Let

M = the event the consumer is a man

W = the event the consumer is a woman

B = the event the consumer preferred plain bottled water

S = the event the consumer preferred sports drink

- What is the probability a person in the study preferred plain bottled water?
- What is the probability a person in the study preferred a sports drink?
- What are the conditional probabilities $P(M | S)$ and $P(W | S)$?
- What are the joint probabilities $P(M \cap S)$ and $P(W \cap S)$?
- Given a consumer is a man, what is the probability he will prefer a sports drink?
- Given a consumer is a woman, what is the probability she will prefer a sports drink?
- Is preference for a sports drink independent of whether the consumer is a man or a woman? Explain using probability information.

4.5 Bayes' Theorem

In the discussion of conditional probability, we indicated that revising probabilities when new information is obtained is an important phase of probability analysis. Often, we begin the analysis with initial or **prior probability** estimates for specific events of interest. Then, from sources such as a sample, a special report, or a product test, we obtain additional information about the events. Given this new information, we update the prior probability values by calculating revised probabilities, referred to as **posterior probabilities**. **Bayes' theorem** provides a means for making these probability calculations. The steps in this probability revision process are shown in Figure 4.9.

FIGURE 4.9 PROBABILITY REVISION USING BAYES' THEOREM

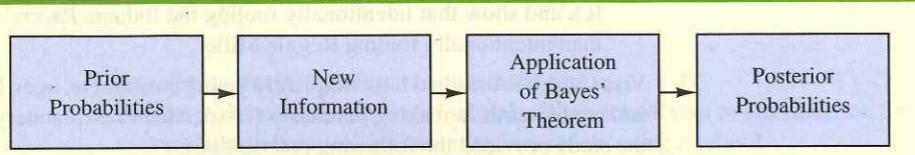


TABLE 4.6 HISTORICAL QUALITY LEVELS OF TWO SUPPLIERS

Supplier	Percentage Good Parts	Percentage Bad Parts
Supplier 1	98	2
Supplier 2	95	5

As an application of Bayes' theorem, consider a manufacturing firm that receives shipments of parts from two different suppliers. Let A_1 denote the event that a part is from supplier 1 and A_2 denote the event that a part is from supplier 2. Currently, 65% of the parts purchased by the company are from supplier 1 and the remaining 35% are from supplier 2. Hence, if a part is selected at random, we would assign the prior probabilities $P(A_1) = .65$ and $P(A_2) = .35$.

The quality of the purchased parts varies with the source of supply. Historical data suggest that the quality ratings of the two suppliers are as shown in Table 4.6. If we let G denote the event that a part is good and B denote the event that a part is bad, the information in Table 4.6 provides the following conditional probability values.

$$P(G | A_1) = .98 \quad P(B | A_1) = .02$$

$$P(G | A_2) = .95 \quad P(B | A_2) = .05$$

The tree diagram in Figure 4.10 depicts the process of the firm receiving a part from one of the two suppliers and then discovering that the part is good or bad as a two-step experiment. We see that four experimental outcomes are possible; two correspond to the part being good and two correspond to the part being bad.

Each of the experimental outcomes is the intersection of two events, so we can use the multiplication rule to compute the probabilities. For instance,

$$P(A_1, G) = P(A_1 \cap G) = P(A_1)P(G | A_1)$$

FIGURE 4.10 TREE DIAGRAM FOR TWO-SUPPLIER EXAMPLE

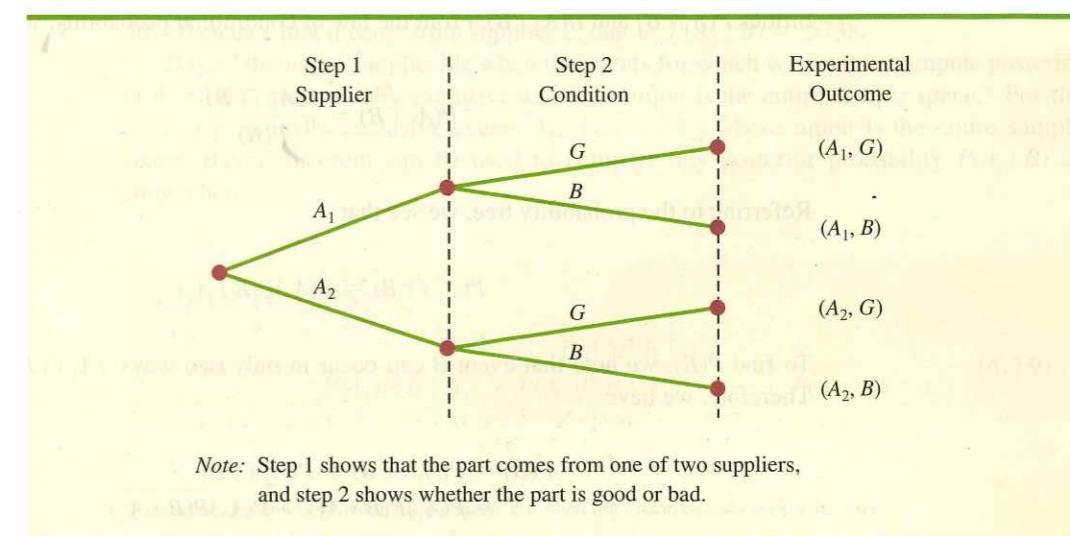
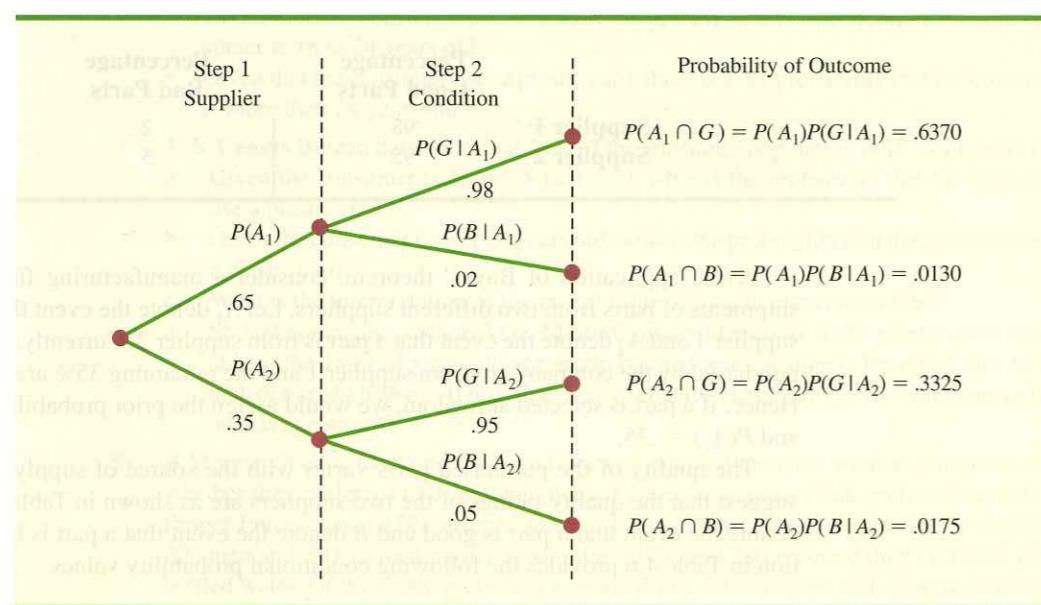


FIGURE 4.11 PROBABILITY TREE FOR TWO-SUPPLIER EXAMPLE

The process of computing these joint probabilities can be depicted in what is called a probability tree (see Figure 4.11). From left to right through the tree, the probabilities for each branch at step 1 are prior probabilities and the probabilities for each branch at step 2 are conditional probabilities. To find the probabilities of each experimental outcome, we simply multiply the probabilities on the branches leading to the outcome. Each of these joint probabilities is shown in Figure 4.11 along with the known probabilities for each branch.

Suppose now that the parts from the two suppliers are used in the firm's manufacturing process and that a machine breaks down because it attempts to process a bad part. Given the information that the part is bad, what is the probability that it came from supplier 1 and what is the probability that it came from supplier 2? With the information in the probability tree (Figure 4.11), Bayes' theorem can be used to answer these questions.

Letting B denote the event that the part is bad, we are looking for the posterior probabilities $P(A_1 | B)$ and $P(A_2 | B)$. From the law of conditional probability, we know that

$$P(A_1 | B) = \frac{P(A_1 \cap B)}{P(B)} \quad (4.14)$$

Referring to the probability tree, we see that

$$P(A_1 \cap B) = P(A_1)P(B | A_1) \quad (4.15)$$

To find $P(B)$, we note that event B can occur in only two ways: $(A_1 \cap B)$ and $(A_2 \cap B)$. Therefore, we have

$$\begin{aligned} P(B) &= P(A_1 \cap B) + P(A_2 \cap B) \\ &= P(A_1)P(B | A_1) + P(A_2)P(B | A_2) \end{aligned} \quad (4.16)$$

Substituting from equations (4.15) and (4.16) into equation (4.14) and writing a similar result for $P(A_2 | B)$, we obtain Bayes' theorem for the case of two events.

BAYES' THEOREM (TWO-EVENT CASE)

$$P(A_1 | B) = \frac{P(A_1)P(B | A_1)}{P(A_1)P(B | A_1) + P(A_2)P(B | A_2)} \quad (4.17)$$

$$P(A_2 | B) = \frac{P(A_2)P(B | A_2)}{P(A_1)P(B | A_1) + P(A_2)P(B | A_2)} \quad (4.18)$$

Using equation (4.17) and the probability values provided in the example, we have

$$\begin{aligned} P(A_1 | B) &= \frac{P(A_1)P(B | A_1)}{P(A_1)P(B | A_1) + P(A_2)P(B | A_2)} \\ &= \frac{(.65)(.02)}{(.65)(.02) + (.35)(.05)} = \frac{.0130}{.0130 + .0175} \\ &= \frac{.0130}{.0305} = .4262 \end{aligned}$$

In addition, using equation (4.18), we find $P(A_2 | B)$.

$$\begin{aligned} P(A_2 | B) &= \frac{(.35)(.05)}{(.65)(.02) + (.35)(.05)} \\ &= \frac{.0175}{.0130 + .0175} = \frac{.0175}{.0305} = .5738 \end{aligned}$$

Note that in this application we started with a probability of .65 that a part selected at random was from supplier 1. However, given information that the part is bad, the probability that the part is from supplier 1 drops to .4262. In fact, if the part is bad, it has better than a 50–50 chance that it came from supplier 2; that is, $P(A_2 | B) = .5738$.

Bayes' theorem is applicable when the events for which we want to compute posterior probabilities are mutually exclusive and their union is the entire sample space.* For the case of n mutually exclusive events A_1, A_2, \dots, A_n , whose union is the entire sample space, Bayes' theorem can be used to compute any posterior probability $P(A_i | B)$ as shown here.

BAYES' THEOREM

$$P(A_i | B) = \frac{P(A_i)P(B | A_i)}{P(A_1)P(B | A_1) + P(A_2)P(B | A_2) + \dots + P(A_n)P(B | A_n)} \quad (4.19)$$

*If the union of events is the entire sample space, the events are said to be *collectively exhaustive*.

With prior probabilities $P(A_1), P(A_2), \dots, P(A_n)$ and the appropriate conditional probabilities $P(B | A_1), P(B | A_2), \dots, P(B | A_n)$, equation (4.19) can be used to compute the posterior probability of the events A_1, A_2, \dots, A_n .

Tabular Approach

A tabular approach is helpful in conducting the Bayes' theorem calculations. Such an approach is shown in Table 4.7 for the parts supplier problem. The computations shown there are done in the following steps.

Step 1. Prepare the following three columns:

Column 1—The mutually exclusive events A_i for which posterior probabilities are desired

Column 2—The prior probabilities $P(A_i)$ for the events

Column 3—The conditional probabilities $P(B | A_i)$ of the new information B given each event

Step 2. In column 4, compute the joint probabilities $P(A_i \cap B)$ for each event and the new information B by using the multiplication law. These joint probabilities are found by multiplying the prior probabilities in column 2 by the corresponding conditional probabilities in column 3; that is, $P(A_i \cap B) = P(A_i)P(B | A_i)$.

Step 3. Sum the joint probabilities in column 4. The sum is the probability of the new information, $P(B)$. Thus we see in Table 4.7 that there is a .0130 probability that the part came from supplier 1 and is bad and a .0175 probability that the part came from supplier 2 and is bad. Because these are the only two ways in which a bad part can be obtained, the sum $.0130 + .0175$ shows an overall probability of .0305 of finding a bad part from the combined shipments of the two suppliers.

Step 4. In column 5, compute the posterior probabilities using the basic relationship of conditional probability.

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

Note that the joint probabilities $P(A_i \cap B)$ are in column 4 and the probability $P(B)$ is the sum of column 4.

TABLE 4.7 TABULAR APPROACH TO BAYES' THEOREM CALCULATIONS FOR THE TWO-SUPPLIER PROBLEM

(1) Events A_i	(2) Prior Probabilities $P(A_i)$	(3) Conditional Probabilities $P(B A_i)$	(4) Joint Probabilities $P(A_i \cap B)$	(5) Posterior Probabilities $P(A_i B)$
A_1	.65	.02	.0130	.0130/.0305 = .4262
A_2	.35	.05	.0175	.0175/.0305 = .5738
	1.00		$P(B) = .0305$	1.0000

NOTES AND COMMENTS

1. Bayes' theorem is used extensively in decision analysis. The prior probabilities are often subjective estimates provided by a decision maker. Sample information is obtained and posterior probabilities are computed for use in choosing the best decision.
2. An event and its complement are mutually exclusive, and their union is the entire sample space. Thus, Bayes' theorem is always applicable for computing posterior probabilities of an event and its complement.

Exercises

Methods

SELF test

39. The prior probabilities for events A_1 and A_2 are $P(A_1) = .40$ and $P(A_2) = .60$. It is also known that $P(A_1 \cap A_2) = 0$. Suppose $P(B | A_1) = .20$ and $P(B | A_2) = .05$.
 - a. Are A_1 and A_2 mutually exclusive? Explain.
 - b. Compute $P(A_1 \cap B)$ and $P(A_2 \cap B)$.
 - c. Compute $P(B)$.
 - d. Apply Bayes' theorem to compute $P(A_1 | B)$ and $P(A_2 | B)$.
40. The prior probabilities for events A_1, A_2 , and A_3 are $P(A_1) = .20, P(A_2) = .50$, and $P(A_3) = .30$. The conditional probabilities of event B given A_1, A_2 , and A_3 are $P(B | A_1) = .50, P(B | A_2) = .40$, and $P(B | A_3) = .30$.
 - a. Compute $P(B \cap A_1), P(B \cap A_2)$, and $P(B \cap A_3)$.
 - b. Apply Bayes' theorem, equation (4.19), to compute the posterior probability $P(A_2 | B)$.
 - c. Use the tabular approach to applying Bayes' theorem to compute $P(A_1 | B), P(A_2 | B)$, and $P(A_3 | B)$.

Applications

41. A consulting firm submitted a bid for a large research project. The firm's management initially felt they had a 50–50 chance of getting the project. However, the agency to which the bid was submitted subsequently requested additional information on the bid. Past experience indicates that for 75% of the successful bids and 40% of the unsuccessful bids the agency requested additional information.
 - a. What is the prior probability of the bid being successful (that is, prior to the request for additional information)?
 - b. What is the conditional probability of a request for additional information given that the bid will ultimately be successful?
 - c. Compute the posterior probability that the bid will be successful given a request for additional information.
42. A local bank reviewed its credit card policy with the intention of recalling some of its credit cards. In the past approximately 5% of cardholders defaulted, leaving the bank unable to collect the outstanding balance. Hence, management established a prior probability of .05 that any particular cardholder will default. The bank also found that the probability of missing a monthly payment is .20 for customers who do not default. Of course, the probability of missing a monthly payment for those who default is 1.
 - a. Given that a customer missed one or more monthly payments, compute the posterior probability that the customer will default.
 - b. The bank would like to recall its card if the probability that a customer will default is greater than .20. Should the bank recall its card if the customer misses a monthly payment? Why or why not?

SELF test

43. Small cars get better gas mileage, but they are not as safe as bigger cars. Small cars accounted for 18% of the vehicles on the road, but accidents involving small cars led to 11,898 fatalities during a recent year (*Reader's Digest*, May 2000). Assume the probability a small car is involved in an accident is .18. The probability of an accident involving a small car leading to a fatality is .128 and the probability of an accident not involving a small car leading to a fatality is .05. Suppose you learn of an accident involving a fatality. What is the probability a small car was involved? Assume that the likelihood of getting into an accident is independent of car size.
44. The American Council of Education reported that 47% of college freshmen earn a degree and graduate within five years (Associated Press, May 6, 2002). Assume that graduation records show women make up 50% of the students who graduated within five years, but only 45% of the students who did not graduate within five years. The students who had not graduated within five years either dropped out or were still working on their degrees.
- Let A_1 = the student graduated within five years
 - A_2 = the student did not graduate within five years
 - W = the student is a female student
- Using the given information, what are the values for $P(A_1)$, $P(A_2)$, $P(W | A_1)$, and $P(W | A_2)$?
- What is the probability that a female student will graduate within five years?
 - What is the probability that a male student will graduate within five years?
 - Given the preceding results, what are the percentage of women and the percentage of men in the entering freshman class?
45. In an article about investment growth, *Money* magazine reported that drug stocks show powerful long-term trends and offer investors unparalleled potential for strong and steady gains. The federal Health Care Financing Administration supports this conclusion through its forecast that annual prescription drug expenditures will reach \$366 billion by 2010, up from \$117 billion in 2000. Many individuals age 65 and older rely heavily on prescription drugs. For this group, 82% take prescription drugs regularly, 55% take three or more prescriptions regularly, and 40% currently use five or more prescriptions. In contrast, 49% of people under age 65 take prescriptions regularly, with 37% taking three or more prescriptions regularly and 28% using five or more prescriptions (*Money*, September 2001). The U.S. Census Bureau reports that of the 281,421,906 people in the United States, 34,991,753 are age 65 years and older (U.S. Census Bureau, *Census 2000*).
- Compute the probability that a person in the United States is age 65 or older.
 - Compute the probability that a person takes prescription drugs regularly.
 - Compute the probability that a person is age 65 or older and takes five or more prescriptions.
 - Given a person uses five or more prescriptions, compute the probability that the person is age 65 or older.

Summary

In this chapter we introduced basic probability concepts and illustrated how probability analysis can be used to provide helpful information for decision making. We described how probability can be interpreted as a numerical measure of the likelihood that an event will occur. In addition, we saw that the probability of an event can be computed either by summing the probabilities of the experimental outcomes (sample points) comprising the event or by using the relationships established by the addition, conditional probability, and multiplication laws of probability. For cases in which additional information is available, we showed how Bayes' theorem can be used to obtain revised or posterior probabilities.

Glossary

Probability A numerical measure of the likelihood that an event will occur.
Experiment A process that generates well-defined outcomes.

Sample space The set of all experimental outcomes.

Sample point An element of the sample space. A sample point represents an experimental outcome.

Tree diagram A graphical representation that helps in visualizing a multiple-step experiment.

Basic requirements for assigning probabilities Two requirements that restrict the manner in which probability assignments can be made: (1) for each experimental outcome E_i we must have $0 \leq P(E_i) \leq 1$; (2) considering all experimental outcomes, we must have $P(E_1) + P(E_2) + \dots + P(E_n) = 1.0$.

Classical method A method of assigning probabilities that is appropriate when all the experimental outcomes are equally likely.

Relative frequency method A method of assigning probabilities that is appropriate when data are available to estimate the proportion of the time the experimental outcome will occur if the experiment is repeated a large number of times.

Subjective method A method of assigning probabilities on the basis of judgment.

Event A collection of sample points.

Complement of A The event consisting of all sample points that are not in A .

Venn diagram A graphical representation for showing symbolically the sample space and operations involving events in which the sample space is represented by a rectangle and events are represented as circles within the sample space.

Union of A and B The event containing all sample points belonging to A or B or both. The union is denoted $A \cup B$.

Intersection of A and B The event containing the sample points belonging to both A and B . The intersection is denoted $A \cap B$.

Addition law A probability law used to compute the probability of the union of two events. It is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. For mutually exclusive events, $P(A \cap B) = 0$; in this case the addition law reduces to $P(A \cup B) = P(A) + P(B)$.

Mutually exclusive events Events that have no sample points in common; that is, $A \cap B$ is empty and $P(A \cap B) = 0$.

Conditional probability The probability of an event given that another event already occurred. The conditional probability of A given B is $P(A | B) = P(A \cap B)/P(B)$.

Joint probability The probability of two events both occurring; that is, the probability of the intersection of two events.

Marginal probability The values in the margins of a joint probability table that provide the probabilities of each event separately.

Independent events Two events A and B where $P(A | B) = P(A)$ or $P(B | A) = P(B)$; that is, the events have no influence on each other.

Multiplication law A probability law used to compute the probability of the intersection of two events. It is $P(A \cap B) = P(B)P(A | B)$ or $P(A \cap B) = P(A)P(B | A)$. For independent events it reduces to $P(A \cap B) = P(A)P(B)$.

Prior probabilities Initial estimates of the probabilities of events.

Posterior probabilities Revised probabilities of events based on additional information.

Bayes' theorem A method used to compute posterior probabilities.

Key Formulas

Counting Rule for Combinations

$$C_n^N = \binom{N}{n} = \frac{N!}{n!(N-n)!} \quad (4.1)$$

Counting Rule for Permutations

$$P_n^N = n! \binom{N}{n} = \frac{N!}{(N-n)!} \quad (4.2)$$

Computing Probability Using the Complement

$$P(A) = 1 - P(A^c) \quad (4.5)$$

Addition Law

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (4.6)$$

Conditional Probability

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad (4.7)$$

$$P(B | A) = \frac{P(A \cap B)}{P(A)} \quad (4.8)$$

Multiplication Law

$$P(A \cap B) = P(B)P(A | B) \quad (4.11)$$

$$P(A \cap B) = P(A)P(B | A) \quad (4.12)$$

Multiplication Law for Independent Events

$$P(A \cap B) = P(A)P(B) \quad (4.13)$$

Bayes' Theorem

$$P(A_i | B) = \frac{P(A_i)P(B | A_i)}{P(A_1)P(B | A_1) + P(A_2)P(B | A_2) + \dots + P(A_n)P(B | A_n)} \quad (4.19)$$

Supplementary Exercises

46. Refer to the data from the new-matriculants survey in exercise 48.
- Given that a person applied to more than one school, what is the probability that the person is 24–26 years old?
 - Given that a person is in the 36-and-over age group, what is the probability that the person applied to more than one school?
 - What is the probability that a person is 24–26 years old or applied to more than one school?
 - Suppose a person is known to have applied to only one school. What is the probability that the person is 31 or more years old?
 - Is the number of schools applied to independent of age? Explain.
47. Companies that do business over the Internet can often obtain probability information about Web site visitors from previous Web sites visited. The article “Internet Marketing” (*Interfaces*, March/April 2001) described how clickstream data on Web sites visited could be used in conjunction with a Bayesian updating scheme to determine the gender of a Web site visitor. Par Fore created a Web site to market golf equipment and apparel. Management would like a certain offer to appear for female visitors and a different offer to appear for male visitors. From a sample of past Web site visits, management learned that 60% of the visitors to ParFore.com are male and 40% are female.
- What is the prior probability that the next visitor to the Web site will be female?
 - Suppose you know that the current visitor to ParFore.com previously visited the Dillard’s Web site, and that women are three times as likely to visit the Dillard’s Web

site as men. What is the revised probability that the current visitor to ParFore.com is female? Should you display the offer that appeals more to female visitors or the one that appeals more to male visitors?

48. An MBA new-matriculants survey provided the following data for 2018 students.

		Applied to More Than One School	
		Yes	No
Age Group	23 and under	207	201
	24–26	299	379
	27–30	185	268
	31–35	66	193
	36 and over	51	169

- For a randomly selected MBA student, prepare a joint probability table for the experiment consisting of observing the student’s age and whether the student applied to one or more schools.
 - What is the probability that a randomly selected applicant is 23 or under?
 - What is the probability that a randomly selected applicant is older than 26?
 - What is the probability that a randomly selected applicant applied to more than one school?
49. An oil company purchased an option on land in Alaska. Preliminary geologic studies assigned the following prior probabilities.

$$P(\text{high-quality oil}) = .50$$

$$P(\text{medium-quality oil}) = .20$$

$$P(\text{no oil}) = .30$$

- What is the probability of finding oil?
- After 200 feet of drilling on the first well, a soil test is taken. The probabilities of finding the particular type of soil identified by the test follow.

$$P(\text{soil} | \text{high-quality oil}) = .20$$

$$P(\text{soil} | \text{medium-quality oil}) = .80$$

$$P(\text{soil} | \text{no oil}) = .20$$

How should the firm interpret the soil test? What are the revised probabilities, and what is the new probability of finding oil?

50. The following crosstabulation shows household income by educational level of the head of household (*Statistical Abstract of the United States: 2002*).

Education Level	Household Income (\$1000s)					Total
	Under 25	25.0–49.9	50.0–74.9	75.0–99.9	100 or more	
Not H.S. Graduate	9,285	4,093	1,589	541	354	15,862
H.S. Graduate	10,150	9,821	6,050	2,737	2,028	30,786
Some College	6,011	8,221	5,813	3,215	3,120	26,380
Bachelor's Degree	2,138	3,985	3,952	2,698	4,748	17,521
Beyond Bach. Deg.	813	1,497	1,815	1,589	3,765	9,479
Total	28,397	27,617	19,219	10,780	14,015	100,028

- Develop a joint probability table.
- What is the probability of a head of household not being a high school graduate?

- c. What is the probability of a head of household having a bachelor's degree or more education?
 - d. What is the probability of a household headed by someone with a bachelor's degree earning \$100,000 or more?
 - e. What is the probability of a household having income below \$25,000?
 - f. What is the probability of a household headed by someone with a bachelor's degree earning less than \$25,000?
 - g. Is household income independent of educational level?
51. A survey conducted by the Pew Internet & American Life Project showed that 8% of Internet users age 18 and older report keeping a blog. Referring to the 18–29 age group as young adults, the survey results showed that for bloggers 54% are young adults and for non-bloggers 24% are young adults (Pew Internet & American Life Project, July 19, 2006).
- a. Develop a joint probability table for these data with two rows (bloggers vs. non-bloggers) and two columns (young adults vs. older adults).
 - b. What is the probability that an Internet user is a young adult?
 - c. What is the probability that an Internet user keeps a blog and is a young adult?
 - d. Suppose that in a follow-up phone survey we contact a respondent who is 24 years old. What is the probability that this respondent keeps a blog?
52. A telephone survey to determine viewer response to a new television show obtained the following data.
- | Rating | Frequency |
|---------------|-----------|
| Poor | 4 |
| Below average | 8 |
| Average | 11 |
| Above average | 14 |
| Excellent | 13 |
- a. What is the probability that a randomly selected viewer will rate the new show as average or better?
 - b. What is the probability that a randomly selected viewer will rate the new show below average or worse?
53. A company studied the number of lost-time accidents occurring at its Brownsville, Texas, plant. Historical records show that 6% of the employees suffered lost-time accidents last year. Management believes that a special safety program will reduce such accidents to 5% during the current year. In addition, it estimates that 15% of employees who had lost-time accidents last year will experience a lost-time accident during the current year.
- a. What percentage of the employees will experience lost-time accidents in both years?
 - b. What percentage of the employees will suffer at least one lost-time accident over the two-year period?
54. A study of 31,000 hospital admissions in New York State found that 4% of the admissions led to treatment-caused injuries. One-seventh of these treatment-caused injuries resulted in death, and one-fourth were caused by negligence. Malpractice claims were filed in one out of 7.5 cases involving negligence, and payments were made in one out of every two claims.
- a. What is the probability a person admitted to the hospital will suffer a treatment-caused injury due to negligence?
 - b. What is the probability a person admitted to the hospital will die from a treatment-caused injury?
 - c. In the case of a negligent treatment-caused injury, what is the probability a malpractice claim will be paid?

55. Cooper Realty is a small real estate company located in Albany, New York, specializing primarily in residential listings. They recently became interested in determining the likelihood of one of their listings being sold within a certain number of days. An analysis of company sales of 800 homes in previous years produced the following data.

Initial Asking Price	Days Listed Until Sold			Total
	Under 30	31–90	Over 90	
Under \$150,000	50	40	10	100
\$150,000–\$199,999	20	150	80	250
\$200,000–\$250,000	20	280	100	400
Over \$250,000	10	30	10	50
	Total	100	500	200
				800

- a. If A is defined as the event that a home is listed for more than 90 days before being sold, estimate the probability of A .
 - b. If B is defined as the event that the initial asking price is under \$150,000, estimate the probability of B .
 - c. What is the probability of $A \cap B$?
 - d. Assuming that a contract was just signed to list a home with an initial asking price of less than \$150,000, what is the probability that the home will take Cooper Realty more than 90 days to sell?
 - e. Are events A and B independent?
56. In early 2003, President Bush proposed eliminating the taxation of dividends to shareholders on the grounds that it was double taxation. Corporations pay taxes on the earnings that are later paid out in dividends. In a poll of 671 Americans, TechnoMetrica Market Intelligence found that 47% favored the proposal, 44% opposed it, and 9% were not sure (*Investor's Business Daily*, January 13, 2003). In looking at the responses across party lines the poll showed that 29% of Democrats were in favor, 64% of Republicans were in favor, and 48% of Independents were in favor.
- a. How many of those polled favored elimination of the tax on dividends?
 - b. What is the conditional probability in favor of the proposal given the person polled is a Democrat?
 - c. Is party affiliation independent of whether one is in favor of the proposal?
 - d. If we assume people's responses were consistent with their own self-interest, which group do you believe will benefit most from passage of the proposal?
57. A large consumer goods company ran a television advertisement for one of its soap products. On the basis of a survey that was conducted, probabilities were assigned to the following events.

B = individual purchased the product

S = individual recalls seeing the advertisement

$B \cap S$ = individual purchased the product and recalls seeing the advertisement

The probabilities assigned were $P(B) = .20$, $P(S) = .40$, and $P(B \cap S) = .12$.

- a. What is the probability of an individual's purchasing the product given that the individual recalls seeing the advertisement? Does seeing the advertisement increase the probability that the individual will purchase the product? As a decision maker, would you recommend continuing the advertisement (assuming that the cost is reasonable)?

- b. Assume that individuals who do not purchase the company's soap product buy from its competitors. What would be your estimate of the company's market share? Would you expect that continuing the advertisement will increase the company's market share? Why or why not?
- c. The company also tested another advertisement and assigned it values of $P(S) = .30$ and $P(B \cap S) = .10$. What is $P(B | S)$ for this other advertisement? Which advertisement seems to have had the bigger effect on customer purchases?
58. A financial manager made two new investments—one in the oil industry and one in municipal bonds. After a one-year period, each of the investments will be classified as either successful or unsuccessful. Consider the making of the two investments as an experiment.
- How many sample points exist for this experiment?
 - Show a tree diagram and list the sample points.
 - Let O = the event that the oil industry investment is successful and M = the event that the municipal bond investment is successful. List the sample points in O and in M .
 - List the sample points in the union of the events ($O \cup M$).
 - List the sample points in the intersection of the events ($O \cap M$).
 - Are events O and M mutually exclusive? Explain.
59. An IBD/TIPP poll conducted to learn about attitudes toward investment and retirement (*Investor's Business Daily*, May 5, 2000) asked male and female respondents how important they felt level of risk was in choosing a retirement investment. The following joint probability table was constructed from the data provided. "Important" means the respondent said level of risk was either important or very important.

	Male	Female	Total
Important	.22	.27	.49
Not Important	.28	.23	.51
Total	.50	.50	1.00

- What is the probability a survey respondent will say level of risk is important?
 - What is the probability a male respondent will say level of risk is important?
 - What is the probability a female respondent will say level of risk is important?
 - Is the level of risk independent of the gender of the respondent? Why or why not?
 - Do male and female attitudes toward risk differ?
60. In a *BusinessWeek/Harris Poll*, 1035 adults were asked about their attitudes toward business (*BusinessWeek*, September 11, 2000). One question asked: "How would you rate large U.S. companies on making good products and competing in a global environment?" The responses were: excellent—18%, pretty good—50%, only fair—26%, poor—5%, and don't know/no answer—1%.
- What is the probability that a respondent rated U.S. companies pretty good or excellent?
 - How many respondents rated U.S. companies poor?
 - How many respondents did not know or did not answer?

Case Problem Hamilton County Judges

Hamilton County judges try thousands of cases per year. In an overwhelming majority of the cases disposed, the verdict stands as rendered. However, some cases are appealed, and of those appealed, some of the cases are reversed. Kristen DelGuzzi of *The Cincinnati Enquirer* conducted a study of cases handled by Hamilton County judges over a three-year period. Shown in Table 4.8 are the results for 182,908 cases handled (disposed) by 38 judges

TABLE 4.8 TOTAL CASES DISPOSED, APPEALED, AND REVERSED IN HAMILTON COUNTY COURTS

Common Pleas Court			
Judge	Total Cases Disposed	Appealed Cases	Reversed Cases
Fred Cartolano	3,037	137	12
Thomas Crush	3,372	119	10
Patrick Dinkelacker	1,258	44	8
Timothy Hogan	1,954	60	7
Robert Kraft	3,138	127	7
William Mathews	2,264	91	18
William Morrissey	3,032	121	22
Norbert Nadel	2,959	131	20
Arthur Ney, Jr.	3,219	125	14
Richard Niehaus	3,353	137	16
Thomas Nurre	3,000	121	6
John O'Connor	2,969	129	12
Robert Ruehlman	3,205	145	18
J. Howard Sundermann	955	60	10
Ann Marie Tracey	3,141	127	13
Ralph Winkler	3,089	88	6
Total	43,945	1762	199

Domestic Relations Court			
Judge	Total Cases Disposed	Appealed Cases	Reversed Cases
Penelope Cunningham	2,729	7	1
Patrick Dinkelacker	6,001	19	4
Deborah Gaines	8,799	48	9
Ronald Panioto	12,970	32	3
Total	30,499	106	17

Municipal Court			
Judge	Total Cases Disposed	Appealed Cases	Reversed Cases
Mike Allen	6,149	43	4
Nadine Allen	7,812	34	6
Timothy Black	7,954	41	6
David Davis	7,736	43	5
Leslie Isaiah Gaines	5,282	35	13
Karla Grady	5,253	6	0
Deidra Hair	2,532	5	0
Dennis Helmick	7,900	29	5
Timothy Hogan	2,308	13	2
James Patrick Kenney	2,798	6	1
Joseph Luebbers	4,698	25	8
William Mallory	8,277	38	9
Melba Marsh	8,219	34	7
Beth Mattingly	2,971	13	1
Albert Mestemaker	4,975	28	9
Mark Painter	2,239	7	3
Jack Rosen	7,790	41	13
Mark Schweikert	5,403	33	6
David Stockdale	5,371	22	4
John A. West	2,797	4	2
Total	108,464	500	104

in Common Pleas Court, Domestic Relations Court, and Municipal Court. Two of the judges (Dinkelacker and Hogan) did not serve in the same court for the entire three-year period.

The purpose of the newspaper's study was to evaluate the performance of the judges. Appeals are often the result of mistakes made by judges, and the newspaper wanted to know which judges were doing a good job and which were making too many mistakes. You are called in to assist in the data analysis. Use your knowledge of probability and conditional probability to help with the ranking of the judges. You also may be able to analyze the likelihood of appeal and reversal for cases handled by different courts.

Managerial Report

Prepare a report with your rankings of the judges. Also, include an analysis of the likelihood of appeal and case reversal in the three courts. At a minimum, your report should include the following:

1. The probability of cases being appealed and reversed in the three different courts.
2. The probability of a case being appealed for each judge.
3. The probability of a case being reversed for each judge.
4. The probability of reversal given an appeal for each judge.
5. Rank the judges within each court. State the criteria you used and provide a rationale for your choice.

CHAPTER 5



Discrete Probability Distributions

CONTENTS

STATISTICS IN PRACTICE: CITIBANK

- 5.1 RANDOM VARIABLES**
Discrete Random Variables
Continuous Random Variables
- 5.2 DISCRETE PROBABILITY DISTRIBUTIONS**
- 5.3 EXPECTED VALUE AND VARIANCE**
Expected Value
Variance
- 5.4 BINOMIAL PROBABILITY DISTRIBUTION**
A Binomial Experiment

- Martin Clothing Store Problem
Using Tables of Binomial Probabilities
Expected Value and Variance for the Binomial Distribution
- 5.5 POISSON PROBABILITY DISTRIBUTION**
An Example Involving Time Intervals
An Example Involving Length or Distance Intervals
- 5.6 HYPERGEOMETRIC PROBABILITY DISTRIBUTION**