CHAPTER 4

Probability

LEARNING OBJECTIVES

The main objective of Chapter 4 is to help you understand the basic principles of probability, thereby enabling you to:

1. Describe what probability is and when one would use it
2. Differentiate among three methods of assigning probabilities: the classical method, relative frequency of occurrence, and subjective probability
3. Deconstruct the elements of probability by defining experiments, sample spaces, and events, classifying events as mutually exclusive, collectively exhaustive, complementary, or independent, and counting possibilities
4. Compare marginal, union, joint, and conditional probabilities by defining each one.
5. Calculate probabilities using the general law of addition, along with a probability matrix, the complement of a union, or the special law of addition if necessary
6. Calculate joint probabilities of both independent and dependent events using the general and special laws of multiplication
7. Calculate conditional probabilities with various forms of the law of conditional probability, and use them to determine if two events are independent.
8. Calculate conditional probabilities using Bayes' rule

AAGAMIA/Iconica/Getty Images

Equity of the Sexes in the Workplace

The Civil Rights Act was signed into law in the United States in 1964 by President Lyndon Johnson. This law, which was amended in 1972, resulted in several "titles" that addressed discrimination in American society at various levels. One is Title VII, which pertains specifically to employment discrimination. It applies to all employers with more than 15 employees, along with other institutions. One of the provisions of Title VII makes it illegal to refuse to hire a person on the basis of the person's sex.

Today, company hiring procedures must be within the perview and framework of the Equal Employment Opportunity Commission (EEOC) guidelines and Title VII. How does a company defend its hiring practices or know when they are within acceptable bounds? How can individuals or groups who feel they have been the victims of illegal hiring practices "prove" their case? How can a group demonstrate that they have been "adversely impacted" by a company's discriminatory hiring practices?

Statistics are widely used in employment discrimination actions and by companies in attempting to meet EEOC guidelines. Substantial quantities of human resources data are logged and analyzed on a daily basis.

Client Company Human Resource Data by Sex

	Sex		
Type of Position	**Male**	**Female**	**Total**
Managerial	8	3	11
Professional	31	13	44
Technical	52	17	69
Clerical	9	22	31
Total	100	55	155

Managerial and Statistical Questions

Assume that a small portion of the human resource data was gathered on a client company.

1. Suppose some legal concern has been expressed that a disproportionate number of managerial people at the client company are men. If a worker is randomly selected from the client company, what is the probability that the worker is a woman? If a managerial person is randomly selected, what is the probability that the person is a woman? What factors might enter into the apparent discrepancy between probabilities?
2. Suppose a special bonus is being given to one person in the technical area this year. If the bonus is randomly awarded, what is the probability that it will go to a woman, given that worker is in the technical area? Is this discrimination against male technical workers? What factors might enter into the awarding of the bonus other than random selection?
3. Suppose that at the annual holiday party the name of an employee of the client company will be drawn randomly to win a trip to Hawaii. What is the probability that a professional person will be the winner?
4. What is the probability that the winner will be either a man or a clerical worker? What is the probability that the winner will be a woman and in management? Suppose the winner is a man. What is the probability that the winner is from the technical group, given that the winner is a man?

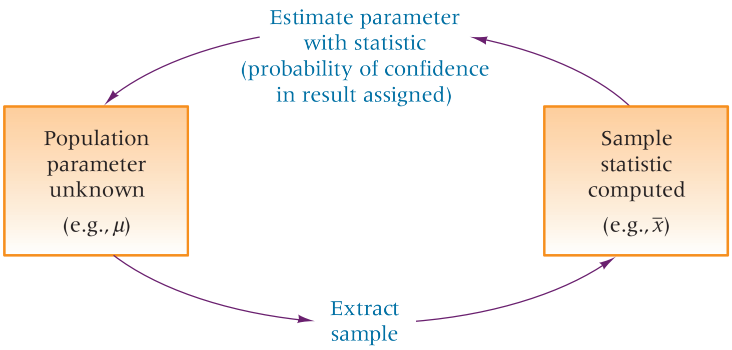
Source: EEOC information adapted from Richard D. Arvey and Robert H. Faley, *Fairness in Selecting Employees,* 2nd ed. Reading, MA: Addison-Wesley Publishing Company, 1992.

In business, most decision making involves uncertainty. For example, an operations manager does not know definitely whether a valve in the plant is going to malfunction or continue to function—or, if it continues, for how long. When should it be replaced? What is the chance that the valve will malfunction within the next week? In the banking industry, what are the new vice president's prospects for successfully turning a department around? The answers to these questions are uncertain.

In the case of a high-rise building, what are the chances that a fire-extinguishing system will work when needed if redundancies are built in? Businesspeople must address these and thousands of similar questions daily. Because most such questions do not have definite answers, the decision making is based on uncertainty. In many of these situations, a probability can be assigned to the likelihood of an outcome. This chapter is about learning how to determine or assign probabilities.

FIGURE 4.1

Probability in the Process of Inferential Statistics



4.1 INTRODUCTION TO PROBABILITY

Chapter 1 discussed the difference between descriptive and inferential statistics. Much statistical analysis is inferential, and probability is the basis for inferential statistics. Recall that inferential statistics involves taking a sample from a population, computing a statistic on the sample, and inferring from the statistic the value of the corresponding parameter of the population. The reason for doing so is that the value of the parameter is unknown. Because it is unknown, the analyst conducts the inferential process under uncertainty. However, by applying rules and laws, the analyst can often assign a probability of obtaining the results. Figure 4.1 depicts this process.

Suppose a quality control inspector selects a random sample of 40 lightbulbs from a population of brand X bulbs and computes the average number of hours of luminance for the sample bulbs. By using techniques discussed later in this text, the specialist estimates the average number of hours of luminance for the *population* of brand X lightbulbs from this sample information. Because the lightbulbs being analyzed are only a sample of the population, the average number of hours of luminance for the 40 bulbs may or may not accurately estimate the average for all bulbs in the population. The results are uncertain. By applying the laws presented in this chapter, the inspector can assign a value of probability to this estimate.

In addition, probabilities are used directly in certain industries and industry applications. For example, the insurance industry uses probabilities in actuarial tables to determine the likelihood of certain outcomes in order to set specific rates and coverages. The gaming industry uses probability values to establish charges and payoffs. One way to determine whether a company's hiring practices meet the government's EEOC guidelines mentioned in the Decision Dilemma is to compare various proportional breakdowns of their employees (by ethnicity, gender, age, etc.) to the proportions in the general population from which the employees are hired. In comparing the company figures with those of the general population, the courts could study the probabilities of a company randomly hiring a certain profile of employees from a given population. In other industries, such as manufacturing and aerospace, it is important to know the life of a mechanized part and the probability that it will malfunction at any given length of time in order to protect the firm from major breakdowns.

4.2 METHODS OF ASSIGNING PROBABILITIES

The three general methods of assigning probabilities are (1) the classical method, (2) the relative frequency of occurrence method, and (3) subjective probabilities.

Classical Method of Assigning Probabilities

When probabilities are assigned based on laws and rules, the method is referred to as the **classical method of assigning probabilities**. This method involves an experiment, which is *a process that produces outcomes*, and an event, *which is an outcome of an experiment.*

When we assign probabilities using the classical method, the probability of an individual event occurring is determined as the ratio of the number of items in a population containing the event (n_e) to the total number of items in the population (N). That is, $P(E) = n_e/N$. For example, if a company has 200 workers and 70 are female, the probability of randomly selecting a female from this company is 70/200 = .35.

CLASSICAL METHOD OF ASSIGNING PROBABILITIES

$$P(E) = \frac{n_e}{N}$$

where

N = total possible number of outcomes of an experiment

n_e = the number of outcomes in which the event occurs out of N outcomes

Suppose, in a particular plant, three machines make a given product. Machine A always produces 40% of the total number of this product. Ten percent of the items produced by machine A are defective. If the finished products are well mixed with regard to which machine produced them and if one of these products is randomly selected, the classical method of assigning probabilities tells us that the probability that the part was produced by machine A and is defective is .04. This probability can be determined even before the part is sampled because with the classical method, the probabilities can be determined **a priori**; that is, *they can be determined prior to the experiment.*

Because n_e can never be greater than N (no more than N outcomes in the population could possibly have attribute e), the highest value of any probability is 1. If the probability of an outcome occurring is 1, the event is certain to occur. The smallest possible probability is 0. If none of the outcomes of the N possibilities has the desired characteristic, e, the probability is $0/N = 0$, and the event is certain not to occur.

RANGE OF POSSIBLE PROBABILITIES

$$0 \le P(E) \le 1$$

Thus, probabilities are nonnegative proper fractions or nonnegative decimal values greater than or equal to 0 and less than or equal to 1.

Probability values can be converted to percentages by multiplying by 100. Meteorologists often report weather probabilities in percentage form. For example, when they forecast a 60% chance of rain for tomorrow, they are saying that the probability of rain tomorrow is .60.

Relative Frequency of Occurrence

The **relative frequency of occurrence method** of assigning probabilities is based on cumulated historical data. With this method, *the probability of an event occurring is equal to the number of times the event has occurred in the past divided by the total number of opportunities for the event to have occurred.*

PROBABILITY BY RELATIVE FREQUENCY OF OCCURRENCE

$$\frac{\text{Number of Times an Event Occurred}}{\text{Total Number of Opportunities for the Event to Occur}}$$

Relative frequency of occurrence is not based on rules or laws but on what has occurred in the past. For example, a company wants to determine the probability that its inspectors are going to reject the next batch of raw materials from a supplier. Data gathered from company record books show that the supplier sent the company 90 batches in the past, and inspectors rejected 10 of them. By the method of relative frequency of occurrence, the probability of the inspectors rejecting the next batch is 10/90, or .11. If the next batch is rejected, the relative frequency of occurrence probability for the subsequent shipment would change to 11/91 = .12.

Subjective Probability

The **subjective method** of *assigning probability is based on the feelings or insights of the person determining the probability*. Subjective probability comes from the person's intuition or reasoning. Although not a scientific approach to probability, the subjective method often is based on the accumulation of knowledge, understanding, and experience stored and processed in the human mind. At times it is merely a guess. At other times, subjective probability can potentially yield accurate probabilities. Subjective probability can be used to capitalize on the background of experienced workers and managers in decision making.

Suppose a director of transportation for an oil company is asked the probability of getting a shipment of oil out of Saudi Arabia to the United States within three weeks. A director who has scheduled many such shipments, has a knowledge of Saudi politics, and has an awareness of current climatological and economic conditions may be able to give an accurate probability that the shipment can be made on time.

Subjective probability also can be a potentially useful way of tapping a person's experience, knowledge, and insight and using them to forecast the occurrence of some event. An experienced airline mechanic can usually assign a meaningful probability that a particular plane will have a certain type of mechanical difficulty. Physicians sometimes assign subjective probabilities to the life expectancy of people who have cancer.

4.3 STRUCTURE OF PROBABILITY

In the study of probability, developing a language of terms and symbols is helpful. The structure of probability provides a common framework within which the topics of probability can be explored.

Experiment

As previously stated, an **experiment** is *a process that produces outcomes*. Examples of business-oriented experiments with outcomes that can be statistically analyzed might include the following.

- Interviewing 20 randomly selected consumers and asking them which brand of appliance they prefer
- Sampling every 200th bottle of ketchup from an assembly line and weighing the contents
- Testing new pharmaceutical drugs on samples of cancer patients and measuring the patients' improvement
- Auditing every 10th account to detect any errors
- Recording the Dow Jones Industrial Average on the first Monday of every month for 10 years

Event

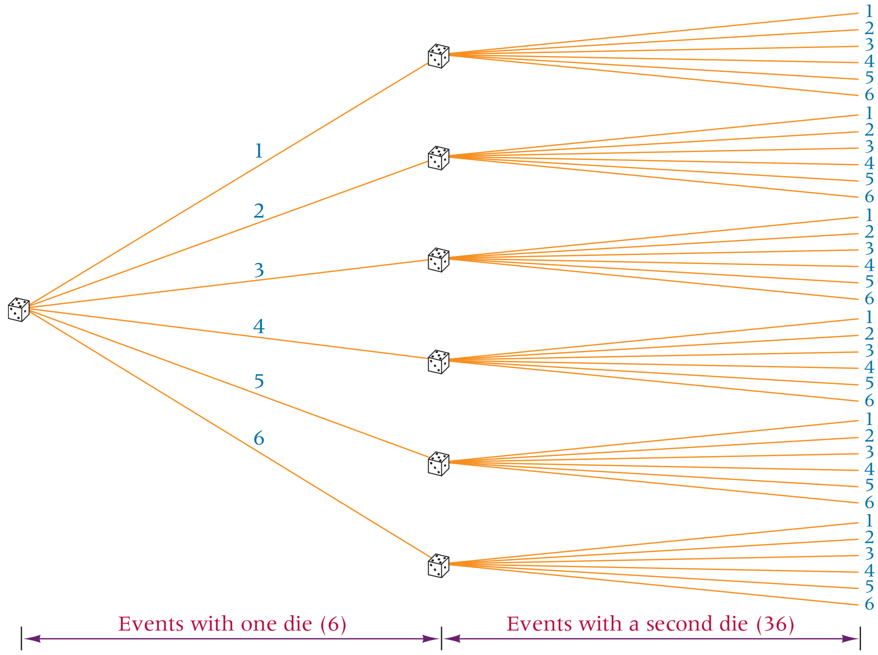
Because an **event** is *an outcome of an experiment*, the experiment defines the possibilities of the event. If the experiment is to sample five bottles coming off a production line, an event could be to get one defective and four good bottles. In an experiment to roll a die, one event could be to roll an even number and another event could be to roll a number greater than two. Events are denoted by uppercase letters; italic capital letters (e.g., A and E_1, E_2, . . .) represent the general or abstract case, and roman capital letters (e.g., H and T for heads and tails) denote specific things and people.

Elementary Events

Events that cannot be decomposed or broken down into other events are called **elementary events**. Elementary events are denoted by lowercase letters (e.g., e_1, e_2, e_3, . . .). Suppose the experiment is to roll a die. The elementary events for this experiment are to roll a 1 or roll a 2 or roll a 3, and so on. Rolling an even number is an event, but it is not an elementary event because the even number can be broken down further into events 2, 4, and 6.

FIGURE 4.2

Possible Outcomes for the Roll of a Pair of Dice



In the experiment of rolling a die, there are six elementary events {1, 2, 3, 4, 5, 6}. Rolling a pair of dice results in 36 possible elementary events (outcomes). For each of the six elementary events possible on the roll of one die, there are six possible elementary events on the roll of the second die, as depicted in the tree diagram in Figure 4.2. Table 4.1 contains a list of these 36 outcomes.

In the experiment of rolling a pair of dice, other events could include outcomes such as two even numbers, a sum of 10, a sum greater than five, and others. However, none of these events is an elementary event because each can be broken down into several of the elementary events displayed in Table 4.1.

Sample Space

A **sample space** is *a complete roster or listing of all elementary events for an experiment*. Table 4.1 is the sample space for the roll of a pair of dice. The sample space for the roll of a single die is {1, 2, 3, 4, 5, 6}.

Sample space can aid in finding probabilities. Suppose an experiment is to roll a pair of dice. What is the probability that the dice will sum to 7? An examination of the sample space shown in Table 4.1 reveals that there are six outcomes in which the dice sum to 7—{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)}—in the total possible 36 elementary events in the sample space. Using this information, we can conclude that the probability of rolling a pair of dice that sum to 7 is 6/36, or .1667. However, using the sample space to determine probabilities is unwieldy and cumbersome when the sample space is large. Hence, statisticians usually use other more effective methods of determining probability.

Unions and Intersections

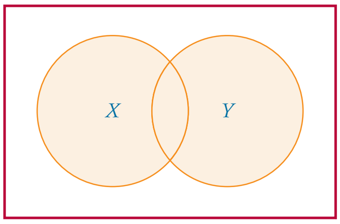
Set notation, the use of braces to group numbers, is used as *a symbolic tool for unions and intersections* in this chapter. The **union** of *X*, *Y* is *formed by combining elements from each of the sets*

TABLE 4.1

All Possible Elementary Events in the Roll of a Pair of Dice (Sample Space)

(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

FIGURE 4.3
A Union



and is denoted $X \cup Y$. An element qualifies for the union of X, Y if it is in either X or Y or in both X and Y. The union expression $X \cup Y$ can be translated to "X or Y." For example, if

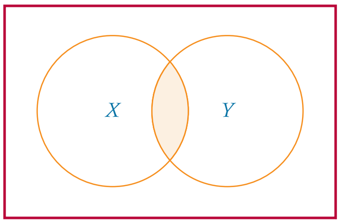
$$X = \{1, 4, 7, 9\} \quad \text{and} \quad Y = \{2, 3, 4, 5, 6\}$$
$$X \cup Y = \{1, 2, 3, 4, 5, 6, 7, 9\}$$

Note that all the values of X and all the values of Y qualify for the union. However, none of the values is listed more than once in the union. In Figure 4.3, the shaded region of the Venn diagram denotes the union.

An intersection is denoted $X \cap Y$. To qualify for intersection, an element must be in both X and Y. The **intersection** *contains the elements common to both sets*. Thus the intersection symbol, $\cap$, is often read as *and*. The intersection of X, Y is referred to as X and Y. For example, if

$$X = \{1, 4, 7, 9\} \quad \text{and} \quad Y = \{2, 3, 4, 5, 6\}$$
$$X \cap Y = \{4\}$$

FIGURE 4.4
An Intersection



Note that only the value 4 is common to both sets X and Y. The intersection is more exclusive than and hence equal to or (usually) smaller than the union. Elements must be characteristic of both X and Y to qualify. In Figure 4.4, the shaded region denotes the intersection.

Mutually Exclusive Events

Two or more events are **mutually exclusive events** if *the occurrence of one event precludes the occurrence of the other event(s)*. This characteristic means that mutually exclusive events cannot occur simultaneously and therefore can have no intersection.

A manufactured part is either defective or okay: The part cannot be both okay and defective at the same time because "okay" and "defective" are mutually exclusive categories. In a sample of the manufactured products, the event of selecting a defective part is mutually exclusive with the event of selecting a nondefective part. Suppose an office building is for sale and two different potential buyers have placed bids on the building. It is not possible for both buyers to purchase the building; therefore, the event of buyer A purchasing the building is mutually exclusive with the event of buyer B purchasing the building. In the toss of a single coin, heads and tails are mutually exclusive events. The person tossing the coin gets either a head or a tail but never both.

The probability of two mutually exclusive events occurring at the same time is zero.

MUTUALLY EXCLUSIVE EVENTS X AND Y

$$P(X \cap Y) = 0$$

Independent Events

Two or more events are **independent events** if *the occurrence or nonoccurrence of one of the events does not affect the occurrence or nonoccurrence of the other event(s)*. Certain experiments, such as rolling dice, yield independent events; each die is independent of the other. Whether a 6 is rolled on the first die has no influence on whether a 6 is rolled on the second die. Coin tosses always are independent of each other. The event of getting a head on the first toss of a coin is independent of getting a head on the second toss. It is generally believed that certain human characteristics are independent of other events. For example, left-handedness is probably independent of the possession of a credit card. Whether a person wears glasses or not is probably independent of the brand of milk preferred.

Many experiments using random selection can produce either independent or nonindependent event, depending on how the experiment is conducted. In these experiments, the outcomes are independent if sampling is done with replacement; that is, after each item is selected and the outcome is determined, the item is restored to the population and the population is shuffled. This way, each draw becomes independent of the previous draw. Suppose an inspector is randomly selecting bolts from a bin that contains 5% defects. If the inspector samples a defective bolt and returns it to the bin, on the second draw there are still 5% defects in the bin regardless of the fact that the first outcome was a defect. If the inspector does not replace the first draw, the second draw is not independent of the first; in this case, fewer than 5% defects remain in the population. Thus the probability of the second outcome is dependent on the first outcome.

If X and Y are independent, the following symbolic notation is used.

INDEPENDENT EVENTS *X* AND *Y*

$$P(X|Y) = P(X) \quad \text{and} \quad P(Y|X) = P(Y)$$

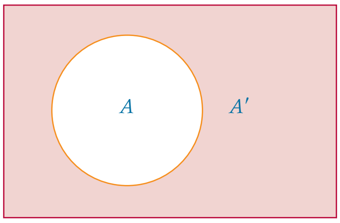
$P(X|Y)$ denotes the probability of X occurring given that Y has occurred. If X and Y are independent, then the probability of X occurring given that Y has occurred is just the probability of X occurring. Knowledge that Y has occurred does not impact the probability of X occurring because X and Y are independent. For example, P (prefers Pepsilperson is right-handed) = P (prefers Pepsi) because a person's handedness is independent of brand preference.

Collectively Exhaustive Events

A list of **collectively exhaustive events** contains *all possible elementary events for an experiment.* Thus, all sample spaces are collectively exhaustive lists. The list of possible outcomes for the tossing of a pair of dice contained in Table 4.1 is a collectively exhaustive list. The sample space for an experiment can be described as a list of events that are mutually exclusive and collectively exhaustive. Sample space events do not overlap or intersect, and the list is complete.

FIGURE 4.5

The Complement of Event *A*



Complementary Events

The **complement** of event A is denoted A', pronounced "not A." All *the elementary events of an experiment not in A comprise its complement.* For example, if in rolling one die, event A is getting an even number, the complement of A is getting an odd number. If event A is getting a 5 on the roll of a die, the complement of A is getting a 1, 2, 3, 4, or 6. The complement of event A contains whatever portion of the sample space that event A does not contain, as the Venn diagram in Figure 4.5 shows.

Using the complement of an event sometimes can be helpful in solving for probabilities because of the following rule.

PROBABILITY OF THE COMPLEMENT OF *A*

$$P(A') = 1 - P(A)$$

Suppose 32% of the employees of a company have a college degree. If an employee is randomly selected from the company, the probability that the person does not have a college degree is $1 - .32 = .68$. Suppose 42% of all parts produced in a plant are molded by machine A and 31% are molded by machine B. If a part is randomly selected, the probability that it was molded by neither machine A nor machine B is $1 - .73 = .27$. (Assume that a part is only molded on one machine.)

Counting the Possibilities

In statistics, a collection of techniques and rules for counting the number of outcomes that can occur for a particular experiment can be used. Some of these rules and techniques can delineate the size of the sample space. Presented here are three of these counting methods.

The *mn* Counting Rule

Suppose a customer decides to buy a certain brand of new car. Options for the car include two different engines, five different paint colors, and three interior packages. If each of these options is available with each of the others, how many different cars could the customer choose from? To determine this number, we can use the ***mn* counting rule**.

THE *mn* COUNTING RULE	For an operation that can be done m ways and a second operation that can be done n ways, the two operations then can occur, in order, in mn ways. This rule can be extended to cases with three or more operations.

Using the mn counting rule, we can determine that the automobile customer has $(2)(5)(3) = 30$ different car combinations of engines, paint colors, and interiors available.

Suppose a scientist wants to set up a research design to study the effects of sex (M, F), marital status (single never married, divorced, married), and economic class (lower, middle, and upper) on the frequency of airline ticket purchases per year. The researcher would set up a design in which 18 different samples are taken to represent all possible groups generated from these customer characteristics.

$$\text{Number of Groups} = (\text{Sex})(\text{Marital Status})(\text{Economic Class}) = (2)(3)(3) = 18 \text{ Groups}$$

Sampling from a Population with Replacement

In the second counting method, sampling n items from a population of size N *with replacement* would provide

$$(N)^n \text{ possibilities}$$

where

N = population size
n = sample size

For example, each time a die, which has six sides, is rolled, the outcomes are independent (with replacement) of the previous roll. If a die is rolled three times in succession, how many different outcomes can occur? That is, what is the size of the sample space for this experiment? The size of the population, N, is 6, the six sides of the die. We are sampling three dice rolls, $n = 3$. The sample space is

$$(N)^n = (6)^3 = 216$$

Suppose in a lottery six numbers are drawn from the digits 0 through 9, with replacement (digits can be reused). How many different groupings of six numbers can be drawn? N is the population of 10 numbers (0 through 9) and n is the sample size, six numbers.

$$(N)^n = (10)^6 = 1{,}000{,}000$$

That is, a million six-digit numbers are available!

Combinations: Sampling from a Population Without Replacement

The third counting method uses **combinations**, sampling n items from a population of size N without replacement provides

$$_NC_n = \binom{N}{n} = \frac{N!}{n!(N-n)!}$$

possibilities.

For example, suppose a small law firm has 16 employees and three are to be selected randomly to represent the company at the annual meeting of the American Bar Association. How many different combinations of lawyers could be sent to the meeting? This situation does not allow sampling with replacement because three *different* lawyers will be selected to go. This problem is solved by using combinations. $N = 16$ and $n = 3$, so

$$_NC_n = {}_{16}C_3 = \frac{16!}{3!13!} = 560$$

A total of 560 combinations of three lawyers could be chosen to represent the firm.

4.3 PROBLEMS

4.1 A supplier shipped a lot of six parts to a company. The lot contained three defective parts. Suppose the customer decided to randomly select two parts and test them for defects. How large a sample space is the customer potentially working with? List the sample space. Using the sample space list, determine the probability that the customer will select a sample with exactly one defect.

4.2 Given X = {1, 3, 5, 7, 8, 9}, Y = {2, 4, 7, 9}, and Z = {1, 2, 3, 4, 7}, solve the following.

a. $X \cup Z =$ _____
b. $X \cap Y =$ _____
c. $X \cap Z =$ _____
d. $X \cup Y \cup Z =$ _____
e. $X \cap Y \cap Z =$ _____
f. $(X \cup Y) \cap Z =$ _____
g. $(Y \cap Z) \cup (X \cap Y) =$ _____
h. X or Y = _____
i. Y and X = _____

4.3 If a population consists of the positive even numbers through 30 and if A = {2, 6, 12, 24}, what is A′?

4.4 A company's customer service 800 telephone system is set up so that the caller has six options. Each of these six options leads to a menu with four options. For each of these four options, three more options are available. For each of these three options, another three options are presented. If a person calls the 800 number for assistance, how many total options are possible?

4.5 A bin contains six parts. Two of the parts are defective and four are acceptable. If three of the six parts are selected from the bin, how large is the sample space? Which counting rule did you use, and why? For this sample space, what is the probability that exactly one of the three sampled parts is defective?

4.6 A company places a seven-digit serial number on each part that is made. Each digit of the serial number can be any number from 0 through 9. Digits can be repeated in the serial number. How many different serial numbers are possible?

4.7 A small company has 20 employees. Six of these employees will be selected randomly to be interviewed as part of an employee satisfaction program. How many different groups of six can be selected?

4.4 MARGINAL, UNION, JOINT, AND CONDITIONAL PROBABILITIES

Four particular types of probability are presented in this chapter. The first type is **marginal probability**. Marginal probability is denoted $P(E)$, where E is some event. A marginal probability is usually *computed by dividing some subtotal by the whole*. An example of marginal probability is the probability that a person owns a Ford car. This probability is computed by dividing the number of Ford owners by the total number of car owners. The probability of a person wearing glasses is also a marginal probability. This probability is computed by dividing the number of people wearing glasses by the total number of people.

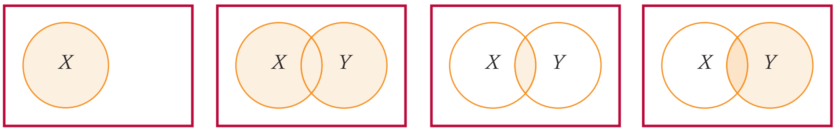
A second type of probability is the union of two events. Union probability is denoted $P(E_1 \cup E_2)$, where E_1 and E_2 are two events. $P(E_1 \cup E_2)$ is the probability that E_1 will occur or that E_2 will occur or that both E_1 and E_2 will occur. An example of union probability is the probability that a person owns a Ford or a Chevrolet. To qualify for the union, the person only has to have at least one of these cars. Another example is the probability of a person wearing glasses or having red hair. All people wearing glasses are included in the union, along with all redheads and all redheads who wear glasses. In a company, the probability that a person is male or a clerical worker is a union probability. A person qualifies for the union by being male or by being a clerical worker or by being both (a male clerical worker).

A third type of probability is the intersection of two events, or joint probability. The joint probability of events E_1 and E_2 occurring is denoted $P(E_1 \cap E_2)$. Sometimes $P(E_1 \cap E_2)$ is read as the probability of E_1 and E_2. To qualify for the intersection, both events must occur. An example of joint probability is the probability of a person owning both a Ford and a Chevrolet. Owning one type of car is not sufficient. A second example of joint probability is the probability that a person is a redhead and wears glasses.

FIGURE 4.6

Marginal, Union, Joint, and Conditional Probabilities

Marginal	Union	Joint	Conditional
$P(X)$	$P(X \cup Y)$	$P(X \cap Y)$	$P(X \mid Y)$
The probability of X occurring	The probability of X or Y occurring	The probability of X and Y occurring	The probability of X occurring given that Y has occurred
Uses total possible outcomes in denominator	Uses total possible outcomes in denominator	Uses total possible outcomes in denominator	Uses subtotal of the possible outcomes in denominator



The fourth type is conditional probability. Conditional probability is denoted $P(E_1 \mid E_2)$. This expression is read: the probability that E_1 will occur given that E_2 is known to have occurred. Conditional probabilities involve knowledge of some prior information. The information that is known or given is written to the right of the vertical line in the probability statement. An example of conditional probability is the probability that a person owns a Chevrolet given that she owns a Ford. This conditional probability is only a measure of the proportion of Ford owners who have a Chevrolet—not the proportion of total car owners who own a Chevrolet. Conditional probabilities are computed by determining the number of items that have an outcome out of some subtotal of the population. In the car owner example, the possibilities are reduced to Ford owners, and then the number of Chevrolet owners out of those Ford owners is determined. Another example of a conditional probability is the probability that a worker in a company is a professional given that he is male. Of the four probability types, only conditional probability does not have the population total as its denominator. Conditional probabilities have a population subtotal in the denominator. Figure 4.6 summarizes these four types of probability.

STATISTICS IN BUSINESS TODAY

Probabilities in the Dry Cleaning Business

According to the International Fabricare Institute, about two-thirds or 67% of all dry cleaning customers are female, and 65% are married. Thirty-seven percent of dry cleaning customers use a cleaner that is within a mile of their home. Do dry cleaning customers care about coupons? Fifty-one percent of dry cleaning customers say that coupons or discounts are important, and in fact, 57% would try another cleaner if a discount were offered. Converting these percentages to proportions, each could be considered to be a marginal probability. For example, if a customer is randomly selected from the dry-cleaning industry, there is a .37 probability that he/she uses a dry cleaner within a mile of his/her home, $P(\leq 1 \text{ mile}) = .37$.

Suppose further analysis shows that 55% of dry-cleaning customers are female and married. Converting this figure to probability results in the joint probability: $P(\text{F} \cap \text{M}) = .55$. Subtracting this value from the .67 who are female, we can determine that 11% of dry cleaning customers are female and not married: $P(\text{F} \cap \text{not M}) = .11$. Suppose 90% of those who say that coupons or discounts are important would try another cleaner if a discount were offered. This can be restated as a conditional probability: $P(\text{try another} \mid \text{coupons important}) = .90$.

Each of the four types of probabilities discussed in this chapter can be applied to the data on consumers in the dry-cleaner industry. Further breakdowns of these statistics using probabilities can offer insights into how to better serve dry-cleaning customers and how to better market dry-cleaning services and products.

4.5 ADDITION LAWS

Video

Several tools are available for use in solving probability problems. These tools include sample space, tree diagrams, the laws of probability, probability matrices, and insight. Because of the individuality and variety of probability problems, some techniques apply more readily in certain situations than in others. No best method is available for solving all probability problems. In some instances, the probability matrix lays out a problem in a readily solvable manner. In other cases, setting up the probability matrix is more difficult than solving the problem in another way. The probability laws almost always can be used to solve probability problems.

Four laws of probability are presented in this chapter: The addition laws, conditional probability, the multiplication laws, and Bayes' rule. The addition laws and the multiplication laws each have a general law and a special law.

The general law of addition is used to find the probability of the union of two events, $P(X \cup Y)$. The expression $P(X \cup Y)$ denotes the probability of X occurring or Y occurring or both X and Y occurring.

GENERAL LAW OF ADDITION

$$P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)$$

where X, Y are events and $(X \cap Y)$ is the intersection of X and Y.

Yankelovich Partners conducted a survey for the American Society of Interior Designers in which workers were asked which changes in office design would increase productivity. Respondents were allowed to answer more than one type of design change. The number one change that 70% of the workers said would increase productivity was reducing noise. In second place was more storage/filing space, selected by 67%. If one of the survey respondents was randomly selected and asked what office design changes would increase worker productivity, what is the probability that this person would select reducing noise *or* more storage/filing space?

Let N represent the event "reducing noise." Let S represent the event "more storage/filing space." The probability of a person responding with *N or S* can be symbolized statistically as a union probability by using the law of addition.

$$P(\text{N} \cup \text{S})$$

To successfully satisfy the search for a person who responds with reducing noise *or* more storage/filing space, we need only find someone who wants *at least one* of those two events. Because 70% of the surveyed people responded that reducing noise would create more productivity, $P(\text{N}) = .70$. In addition, because 67% responded that increased storage space would improve productivity, $P(\text{S}) = .67$. Either of these would satisfy the requirement of the union. Thus, the solution to the problem seems to be

$$P(\text{N} \cup \text{S}) = P(\text{N}) + P(\text{S}) = .70 + .67 = 1.37$$

However, we already established that probabilities cannot be more than 1. What is the problem here? Notice that all people who responded that *both* reducing noise *and* increasing storage space would improve productivity are included in *each* of the marginal probabilities $P(\text{N})$ and $P(\text{S})$. Certainly a respondent who recommends both of these improvements should be included as favoring at least one. However, because they are included in the $P(\text{N})$ *and* the $P(\text{S})$, the people who recommended both improvements are *double counted*. For that reason, the general law of addition subtracts the intersection probability, $P(\text{N} \cap \text{S})$.

In Figure 4.7, Venn diagrams illustrate this discussion. Notice that the intersection area of N and S is double shaded in diagram A, indicating that it has been counted twice. In diagram B, the shading is consistent throughout N and S because the intersection area has been subtracted out. Thus diagram B illustrates the proper application of the general law of addition.

So what is the answer to Yankelovich Partners' union probability question? Suppose 56% of all respondents to the survey had said that *both* noise reduction *and* increased

FIGURE 4.7

Solving for the Union in the Office Productivity Problem

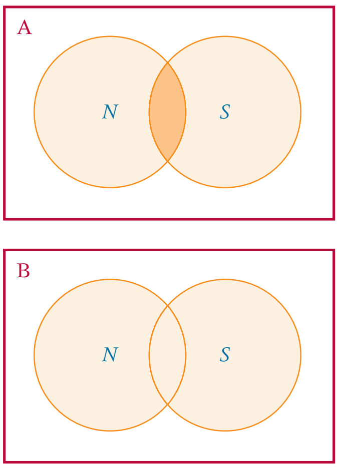


TABLE 4.2

Probability Matrix for the Office Design Problem

		Increase Storage Space	
		Yes	*No*
Noise Reduction	*Yes*		
	No		

storage/filing space would improve productivity: $P(\text{N} \cap \text{S}) = .56$. Then we could use the general law of addition to solve for the probability that a person responds that *either* noise reduction *or* increased storage space would improve productivity.

$$P(\text{N} \cup \text{S}) = P(\text{N}) + P(\text{S}) - P(\text{N} \cap \text{S}) = .70 + .67 - .56 = .81$$

Hence, 81% of the workers surveyed responded that *either* noise reduction *or* increased storage space would improve productivity.

Probability Matrices

In addition to the formulas, another useful tool in solving probability problems is a probability matrix. A **probability matrix** *displays the marginal probabilities and the intersection probabilities of a given problem.* Union probabilities and conditional probabilities must be computed from the matrix. Generally, a probability matrix is constructed as a two-dimensional table with one variable on each side of the table. For example, in the office design problem, noise reduction would be on one side of the table and increased storage space on the other. In this problem, a Yes row and a No row would be created for one variable and a Yes column and a No column would be created for the other variable, as shown in Table 4.2.

Once the matrix is created, we can enter the marginal probabilities. $P(\text{N}) = .70$ is the marginal probability that a person responds yes to noise reduction. This value is placed in the "margin" in the row of Yes to noise reduction, as shown in Table 4.3. If $P(\text{N}) = .70$, then 30% of the people surveyed did not think that noise reduction would increase productivity. Thus, $P(\text{not N}) = 1 - .70 = .30$. This value, also a marginal probability, goes in the row indicated by No under noise reduction. In the column under Yes for increased storage space, the marginal probability $P(\text{S}) = .67$ is recorded. Finally, the marginal probability of No for increased storage space, $P(\text{not S}) = 1 - .67 = .33$, is placed in the No column.

In this probability matrix, all four marginal probabilities are given or can be computed simply by using the probability of a complement rule, $P(\text{not S}) = 1 - P(\text{S})$. The intersection of noise reduction and increased storage space is given as $P(\text{N} \cap \text{S}) = .56$. This value is entered into the probability matrix in the cell under Yes Yes, as shown in Table 4.3. The rest of the matrix can be determined by subtracting the cell values from the marginal probabilities. For example, subtracting .56 from .70 and getting .14 yields the value for the cell under Yes for noise reduction and No for increased storage space. In other words, 14% of all respondents

TABLE 4.3

Probability Matrix for the Office Design Problem

		Increase Storage Space		
		Yes	*No*	
Noise Reduction	*Yes*	.56	.14	.70
	No	.11	.19	.30
		.67	.33	1.00

TABLE 4.4

Yes Row and Yes Column for Probability Matrix of the Office Design Problem

		Increase Storage Space		
		Yes	No	
Noise Reduction	Yes	.56	.14	.70
	No	.11		
		.67		

said that noise reduction would improve productivity but increased storage space would not. Filling out the rest of the matrix results in the probabilities shown in Table 4.3.

Now we can solve the union probability, $P(N \cup S)$, in at least two different ways using the probability matrix. The focus is on the Yes row for noise reduction and the Yes column for increase storage space, as displayed in Table 4.4. The probability of a person suggesting noise reduction *or* increased storage space as a solution for improving productivity, $P(N \cup S)$, can be determined from the probability matrix by adding the marginal probabilities of Yes for noise reduction and Yes for increased storage space and then subtracting the Yes Yes cell, following the pattern of the general law of probabilities.

$$P(N \cup S) = .70 \text{ (from Yes row)} + .67 \text{ (from Yes column)} - .56 \text{ (From Yes Yes cell)} = .81$$

Another way to solve for the union probability from the information displayed in the probability matrix is to sum all cells in any of the Yes rows or columns. Observe the following from Table 4.4.

$$\begin{aligned} P(N \cup S) &= .56 \text{ (from Yes Yes cell)} \\ &\quad + .14 \text{ (from Yes on noise reduction and No on increase storage space)} \\ &\quad + .11 \text{ (from No on noise reduction and Yes on increase storage space)} \\ &= .81 \end{aligned}$$

DEMONSTRATION PROBLEM 4.1

The client company data from the Decision Dilemma reveal that 155 employees worked one of four types of positions. Shown here again is the raw values matrix (also called a contingency table) with the frequency counts for each category and for subtotals and totals containing a breakdown of these employees by type of position and by sex. If an employee of the company is selected randomly, what is the probability that the employee is female or a professional worker?

COMPANY HUMAN RESOURCE DATA

		Sex		
		Male	Female	
Type of Position	Managerial	8	3	11
	Professional	31	13	44
	Technical	52	17	69
	Clerical	9	22	31
		100	55	155

Solution

Let F denote the event of female and P denote the event of professional worker. The question is

$$P(F \cup P) = ?$$

By the general law of addition,

$$P(F \cup P) = P(F) + P(P) - P(F \cap P)$$

Of the 155 employees, 55 are women. Therefore, $P(F) = 55/155 = .355$. The 155 employees include 44 professionals. Therefore, $P(P) = 44/155 = .284$. Because 13 employees are both female and professional, $P(F \cap P) = 13/155 = .084$. The union probability is solved as

$$P(F \cup P) = .355 + .284 - .084 = .555.$$

To solve this probability using a matrix, you can either use the raw values matrix shown previously or convert the raw values matrix to a probability matrix by dividing every value in the matrix by the value of N, 155. The raw value matrix is used in a manner similar to that of the probability matrix. To compute the union probability of selecting a person who is either female or a professional worker from the raw value matrix, add the number of people in the Female column (55) to the number of people in the Professional row (44), then subtract the number of people in the intersection cell of Female and Professional (13). This step yields the value $55 + 44 - 13 = 86$. Dividing this value (86) by the value of N (155) produces the union probability.

$$P(F \cup P) = 86/155 = .555$$

A second way to produce the answer from the raw value matrix is to add all the cells one time that are in either the Female column or the Professional row

$$3 + 13 + 17 + 22 + 31 = 86$$

and then divide by the total number of employees, $N = 155$, which gives

$$P(F \cup P) = 86/155 = .555$$

DEMONSTRATION PROBLEM 4.2

Demonstration Problem

Shown here are the raw values matrix and corresponding probability matrix for the results of a national survey of 200 executives who were asked to identify the geographic locale of their company and their company's industry type. The executives were only allowed to select one locale and one industry type.

RAW VALUES MATRIX

		Geographic Location				
		Northeast D	*Southeast E*	*Midwest F*	*West G*	
Industry Type	*Finance A*	24	10	8	14	56
	Manufacturing B	30	6	22	12	70
	Communications C	28	18	12	16	74
		82	34	42	42	200

By dividing every value of the raw values matrix by the total (200), the corresponding probability matrix (shown at top of next page) can be constructed.

PROBABILITY MATRIX

		Geographic Location				
		Northeast D	Southeast E	Midwest F	West G	
Industry Type	Finance A	.12	.05	.04	.07	.28
	Manufacturing B	.15	.03	.11	.06	.35
	Communications C	.14	.09	.06	.08	.37
		.41	.17	.21	.21	1.00

Suppose a respondent is selected randomly from these data.

a. What is the probability that the respondent is from the Midwest (F)?

b. What is the probability that the respondent is from the communications industry (C) or from the Northeast (D)?

c. What is the probability that the respondent is from the Southeast (E) or from the finance industry (A)?

Solution

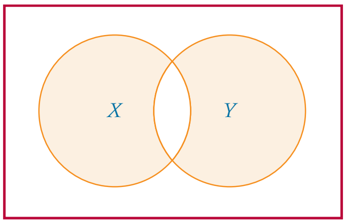
a. $P(\text{Midwest}) = P(\text{F}) = .21$

b. $P(\text{C} \cup \text{D}) = P(\text{C}) + P(\text{D}) - P(\text{C} \cap \text{D}) = .37 + .41 - .14 = .64$

c. $P(\text{E} \cup \text{A}) = P(\text{E}) + P(\text{A}) - P(\text{E} \cap \text{A}) = .17 + .28 - .05 = .40$

FIGURE 4.8

The *X* or *Y* but Not Both Case



In computing the union by using the general law of addition, the intersection probability is subtracted because it is already included in both marginal probabilities. This adjusted probability leaves a union probability that properly includes both marginal values and the intersection value. If the intersection probability is subtracted out a second time, the intersection is removed, leaving the probability of *X or Y* but not *both*.

$$P(X \text{ or } Y \text{ but not both}) = P(X) + P(Y) - P(X \cap Y) - P(X \cap Y)$$
$$= P(X \cup Y) - P(X \cap Y)$$

Figure 4.8 is the Venn diagram for this probability.

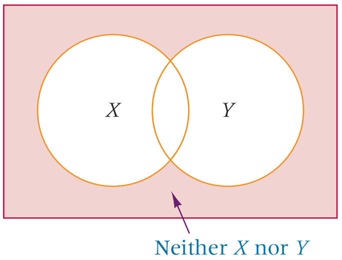
Complement of a Union

The probability of the union of two events *X* and *Y* represents the probability that the outcome is *either X or* it is *Y* or it is *both X* and *Y*. The union includes everything except the possibility that it is neither (*X* or *Y*). Another way to state it is as *neither X nor Y*, which can symbolically be represented as $P(\text{not } X \cap \text{not } Y)$. Because it is the only possible case other than the union of *X* or *Y*, it is the **complement of a union**. Stated more formally,

$$P(\text{neither } X \text{ nor } Y) = P(\text{not } X \cap \text{not } Y) = 1 - P(X \cup Y).$$

FIGURE 4.9

The Complement of a Union: The Neither/Nor Region



Examine the Venn diagram in Figure 4.9. Note that the complement of the union of *X*, *Y* is the shaded area outside the circles. This area represents the neither *X* nor *Y* region.

In the survey about increasing worker productivity by changing the office design discussed earlier, the probability that a randomly selected worker would respond with noise reduction *or* increased storage space was determined to be

$$P(\text{N} \cup \text{S}) = P(\text{N}) + P(\text{S}) - P(\text{N} \cap \text{S}) = .70 + .67 - .56 = .81$$

The probability that a worker would respond with *neither* noise reduction *nor* increased storage space is calculated as the complement of this union.

$$P(\text{neither N nor S}) = P(\text{not N} \cap \text{not S}) = 1 - P(\text{N} \cup \text{S}) = 1 - .81 = .19$$

Thus 19% of the workers selected neither noise reduction nor increased storage space as solutions to increasing productivity. In Table 4.3, this *neither/nor* probability is found in the No No cell of the matrix, .19.

Special Law of Addition

If two events are mutually exclusive, the probability of the union of the two events is the probability of the first event plus the probability of the second event. Because mutually exclusive events do not intersect, nothing has to be subtracted.

SPECIAL LAW OF ADDITION If X, Y are mutually exclusive, $P(X \cup Y) = P(X) + P(Y)$.

The special law of addition is a special case of the general law of addition. In a sense, the general law fits all cases. However, when the events are mutually exclusive, a zero is inserted into the general law formula for the intersection, resulting in the special law formula.

In the survey about improving productivity by changing office design, the respondents were allowed to choose more than one possible office design change. Therefore, it is most likely that virtually none of the change choices were mutually exclusive, and the special law of addition would not apply to that example.

In another survey, however, respondents were allowed to select only one option for their answer, which made the possible options mutually exclusive. In this survey, conducted by Yankelovich Partners for William M. Mercer, Inc., workers were asked what most hinders their productivity and were given only the following selections from which to choose only one answer.

- Lack of direction
- Lack of support
- Too much work
- Inefficient process
- Not enough equipment/supplies
- Low pay/chance to advance

Lack of direction was cited by the most workers (20%), followed by lack of support (18%), too much work (18%), inefficient process (8%), not enough equipment/supplies (7%), low pay/chance to advance (7%), and a variety of other factors added by respondents. If a worker who responded to this survey is selected (or if the survey actually reflects the views of the working public and a worker in general is selected) and that worker is asked which of the given selections most hinders his or her productivity, what is the probability that the worker will respond that it is either too much work or inefficient process?

Let M denote the event "too much work" and I denote the event "inefficient process." The question is:

$$P(\text{M} \cup \text{I}) = ?$$

Because 18% of the survey respondents said "too much work,"

$$P(\text{M}) = .18$$

Because 8% of the survey respondents said "inefficient process,"

$$P(\text{I}) = .08$$

Because it was not possible to select more than one answer,

$$P(\text{M} \cap \text{I}) = .0000$$

Implementing the special law of addition gives

$$P(\text{M} \cup \text{I}) = P(\text{M}) + P(\text{I}) = .18 + .08 = .26$$

DEMONSTRATION PROBLEM 4.3

If a worker is randomly selected from the company described in Demonstration Problem 4.1, what is the probability that the worker is either technical or clerical? What is the probability that the worker is either a professional or a clerical?

Solution

Examine the raw value matrix of the company's human resources data shown in Demonstration Problem 4.1. In many raw value and probability matrices like this one, the rows are nonoverlapping or mutually exclusive, as are the columns. In this matrix, a worker can be classified as being in only one type of position and as either male or female but not both. Thus, the categories of type of position are mutually exclusive, as are the categories of sex, and the special law of addition can be applied to the human resource data to determine the union probabilities.

Let T denote technical, C denote clerical, and P denote professional. The probability that a worker is either technical or clerical is

$$P(\text{T} \cup \text{C}) = P(\text{T}) + P(\text{C}) = \frac{69}{155} + \frac{31}{155} = \frac{100}{155} = .645$$

The probability that a worker is either professional or clerical is

$$P(\text{T} \cup \text{C}) = P(\text{P}) + P(\text{C}) = \frac{44}{155} + \frac{31}{155} = \frac{75}{155} = .484$$

DEMONSTRATION PROBLEM 4.4

Use the data from the matrices in Demonstration Problem 4.2. What is the probability that a randomly selected respondent is from the Southeast or the West?

$$P(\text{E} \cup \text{G}) = ?$$

Solution

Because geographic location is mutually exclusive (the work location is either in the Southeast or in the West but not in both),

$$P(\text{E} \cup \text{G}) = P(\text{E}) + P(\text{G}) = .17 + .21 = .38$$

4.5 PROBLEMS

4.8 Given $P(\text{A}) = .10$, $P(\text{B}) = .12$, $P(\text{C}) = .21$, $P(\text{A} \cap \text{C}) = .05$, and $P(\text{B} \cap \text{C}) = .03$, solve the following.

a. $P(\text{A} \cup \text{C}) =$ _____

b. $P(\text{B} \cup \text{C}) =$ _____

c. If A and B are mutually exclusive, $P(\text{A} \cup \text{B}) =$ _____

4.9 Use the values in the matrix to solve the equations given.

	D	*E*	*F*
A	5	8	12
B	10	6	4
C	8	2	5

a. $P(\text{A} \cup \text{D}) =$ _____

b. $P(\text{E} \cup \text{B}) =$ _____

c. $P(\text{D} \cup \text{E}) =$ _____

d. $P(\text{C} \cup \text{F}) =$ _____

4.10 Use the values in the matrix to solve the equations given.

	E	*F*
A	.10	.03
B	.04	.12
C	.27	.06
D	.31	.07

a. $P(\text{A} \cup \text{F}) =$ _____
b. $P(\text{E} \cup \text{B}) =$ _____
c. $P(\text{B} \cup \text{C}) =$ _____
d. $P(\text{E} \cup \text{F}) =$ _____

4.11 Suppose that 47% of all Americans have flown in an airplane at least once and that 28% of all Americans have ridden on a train at least once. What is the probability that a randomly selected American has either ridden on a train or flown in an airplane? Can this problem be solved? Under what conditions can it be solved? If the problem cannot be solved, what information is needed to make it solvable?

4.12 According to the U.S. Bureau of Labor Statistics, 75% of the women 25 through 49 years of age participate in the labor force. Suppose 78% of the women in that age group are married. Suppose also that 61% of all women 25 through 49 years of age are married and are participating in the labor force.

a. What is the probability that a randomly selected woman in that age group is married or is participating in the labor force?
b. What is the probability that a randomly selected woman in that age group is married or is participating in the labor force but not both?
c. What is the probability that a randomly selected woman in that age group is neither married nor participating in the labor force?

4.13 According to Nielsen Media Research, approximately 67% of all U.S. households with television have cable TV. Seventy-four percent of all U.S. households with television have two or more TV sets. Suppose 55% of all U.S. households with television have cable TV and two or more TV sets. A U.S. household with television is randomly selected.

a. What is the probability that the household has cable TV or two or more TV sets?
b. What is the probability that the household has cable TV or two or more TV sets but not both?
c. What is the probability that the household has neither cable TV nor two or more TV sets?
d. Why does the special law of addition not apply to this problem?

4.14 A survey conducted by the Northwestern University Lindquist-Endicott Report asked 320 companies about the procedures they use in hiring. Only 54% of the responding companies review the applicant's college transcript as part of the hiring process, and only 44% consider faculty references. Assume that these percentages are true for the population of companies in the United States and that 35% of all companies use both the applicant's college transcript and faculty references.

a. What is the probability that a randomly selected company uses either faculty references or college transcript as part of the hiring process?
b. What is the probability that a randomly selected company uses either faculty references or college transcript but not both as part of the hiring process?
c. What is the probability that a randomly selected company uses neither faculty references nor college transcript as part of the hiring process?
d. Construct a probability matrix for this problem and indicate the locations of your answers for parts (a), (b), and (c) on the matrix.

4.6 MULTIPLICATION LAWS

General Law of Multiplication

As stated in Section 4.4, the probability of the intersection of two events $(X \cap Y)$ is called the joint probability. The general law of multiplication is used to find the joint probability.

GENERAL LAW OF MULTIPLICATION

$$P(X \cap Y) = P(X) \cdot P(Y|X) = P(Y) \cdot P(X|Y)$$

The notation $X \cap Y$ means that both *X and Y* must *happen*. The general law of multiplication gives the probability that *both* event *X* and event *Y* will occur at the same time.

According to the U.S. Bureau of Labor Statistics, 46% of the U.S. labor force is female. In addition, 25% of the women in the labor force work part time. What is the probability that a randomly selected member of the U.S. labor force is a woman *and* works part-time? This question is one of joint probability, and the general law of multiplication can be applied to answer it.

Let W denote the event that the member of the labor force is a woman. Let T denote the event that the member is a part-time worker. The question is:

$$P(\text{W} \cap \text{T}) = ?$$

According to the general law of multiplication, this problem can be solved by

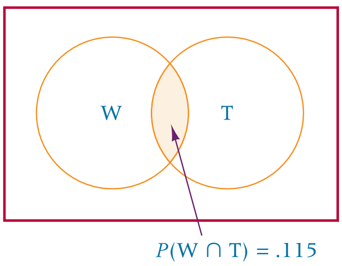
$$P(\text{W} \cap \text{T}) = P(\text{W}) \cdot P(\text{T}|\text{W})$$

Since 46% of the labor force is women, $P(\text{W}) = .46$ $P(\text{T}|\text{W})$ is a conditional probability that can be stated as the probability that a worker is a part-time worker given that the worker is a woman. This condition is what was given in the statement that 25% *of the women in the labor force* work part time. Hence, $P(\text{T}|\text{W}) = .25$. From there it follows that

$$P(\text{W} \cap \text{T}) = P(\text{W}) \cdot P(\text{T}|\text{W}) = (.46)(.25) = .115$$

FIGURE 4.10

Joint Probability that a Woman Is in the Labor Force and Is a Part-Time Worker



It can be stated that 11.5% of the U.S. labor force are women *and* work part-time. The Venn diagram in Figure 4.10 shows these relationships and the joint probability.

Determining joint probabilities from raw value or probability matrices is easy because every cell of these matrices is a joint probability. In fact, some statisticians refer to a probability matrix as a *joint probability table*.

For example, suppose the raw value matrix of the client company data from Demonstration Problem 4.1 and the Decision Dilemma is converted to a probability matrix by dividing by the total number of employees ($N = 155$), resulting in Table 4.5. Each value in the cell of Table 4.5 is an intersection, and the table contains all possible intersections (joint probabilities) for the events of sex and type of position. For example, the probability that a randomly selected worker is male *and* a technical worker, $P(\text{M} \cap \text{T})$, is .335. The probability that a

TABLE 4.5

Probability Matrix of Company Human Resource Data

		Sex		
		Male	*Female*	
Type of Position	*Managerial*	.052	.019	.071
	Professional	.200	.084	.284
	Technical	.335	.110	.445
	Clerical	.058	.142	.200
		.645	.355	1.000

randomly selected worker is female *and* a professional worker, $P(F \cap T)$, is .084. Once a probability matrix is constructed for a problem, usually the easiest way to solve for the joint probability is to find the appropriate cell in the matrix and select the answer. However, sometimes because of what is given in a problem, using the formula is easier than constructing the matrix.

DEMONSTRATION PROBLEM 4.5

A company has 140 employees, of which 30 are supervisors. Eighty of the employees are married, and 20% of the married employees are supervisors. If a company employee is randomly selected, what is the probability that the employee is married and is a supervisor?

Solution

Let M denote married and S denote supervisor. The question is:

$$P(M \cap S) = ?$$

First, calculate the marginal probability.

$$P(M) = \frac{80}{140} = .5714$$

Then, note that 20% of the married employees are supervisors, which is the conditional probability, $P(S|M) = .20$. Finally, applying the general law of multiplication gives

$$P(M \cap S) = P(M) \cdot P(S|M) = (.5714)(.20) = .1143$$

Hence, 11.43% of the 140 employees are married and are supervisors.

DEMONSTRATION PROBLEM 4.6

From the data obtained from the interviews of 200 executives in Demonstration Problem 4.2, find:

a. $P(B \cap E)$
b. $P(G \cap A)$
c. $P(B \cap C)$

PROBABILITY MATRIX

		Geographic Location				
		Northeast D	*Southeast E*	*Midwest F*	*West G*	
Industry Type	*Finance A*	.12	.05	.04	.07	.28
	Manufacturing B	.15	.03	.11	.06	.35
	Communications C	.14	.09	.06	.08	.37
		.41	.17	.21	.21	1.00

Solution

a. From the cell of the probability matrix, $P(B \cap E) = .03$. To solve by the formula, $P(B \cap E) = P(B) \cdot P(E|B)$, first find $P(B)$:

$$P(B) = .35$$

The probability of E occurring given that B has occurred, $P(E|B)$, can be determined from the probability matrix as $P(E|B) = .03/.35$. Therefore,

$$P(B \cap E) = P(B) \cdot P(E|B) = (.35)\left(\frac{.03}{.35}\right) = .03$$

Although the formula works, finding the joint probability in the cell of the probability matrix is faster than using the formula.

An alternative formula is $P(B \cap E) = P(E) \cdot P(B|E)$, but $P(E) = .17$. Then $P(B|E)$ means the probability of B if E is given. There are .17 Es in the probability matrix and .03 Bs in these Es. Hence,

$$P(B|E) = \frac{.03}{.17} \text{ and } P(B \cap E) = P(E) \cdot P(B|E) = (.17)\left(\frac{.03}{.17}\right) = .03$$

b. To obtain $P(G \cap A)$, find the intersecting cell of G and A in the probability matrix, .07, or use one of the following formulas:

$$P(G \cap A) = P(G) \cdot P(A|G) = (.21)\left(\frac{.07}{.21}\right) = .07$$

or

$$P(G \cap A) = P(A) \cdot P(G|A) = (.28)\left(\frac{.07}{.28}\right) = .07$$

c. The probability $P(B \cap C)$ means that one respondent would have to work both in the manufacturing industry and the communications industry. The survey used to gather data from the 200 executives, however, requested that each respondent specify only one industry type for his or her company. The matrix shows no intersection for these two events. Thus B and C are mutually exclusive. None of the respondents is in both manufacturing and communications. Hence,

$$P(B \cap C) = .0$$

Special Law of Multiplication

If events X and Y are independent, a special law of multiplication can be used to find the intersection of X and Y. This special law utilizes the fact that when two events X, Y are independent, $P(X|Y) = P(X)$ and $P(Y|X) = P(Y)$. Thus, the general law of multiplication, $P(X \cap Y) = P(X) \cdot P(Y|X)$, becomes $P(X \cap Y) = P(X) \cdot P(Y)$ when X and Y are independent.

SPECIAL LAW OF MULTIPLICATION

If X, Y are independent, $P(X \cap Y) = P(X) \cdot P(Y)$

A study released by Bruskin-Goldring Research for SEIKO found that 28% of American adults believe that the automated teller has had a most significant impact on everyday life. Another study by David Michaelson & Associates for Dale Carnegie & Associates examined employee views on team spirit in the workplace and discovered that 72% of all employees believe that working as a part of a team lowers stress. Are people's views on automated tellers independent of their views on team spirit in the workplace? If they are independent, then the probability of a person being randomly selected who believes that the automated teller has had a most significant impact on everyday life *and* that working as part of a team lowers stress is found as follows. Let A denote automated teller and S denote teamwork lowers stress.

$$P(A) = .28$$
$$P(S) = .72$$
$$P(A \cap S) = P(A) \cdot P(S) = (.28)(.72) = .2016$$

Therefore, 20.16% of the population believes that the automated teller has had a most significant impact on everyday life *and* that working as part of a team lowers stress.

DEMONSTRATION PROBLEM 4.7

A manufacturing firm produces pads of bound paper. Three percent of all paper pads produced are improperly bound. An inspector randomly samples two pads of paper, one at a time. Because a large number of pads are being produced during the inspection, the sampling being done, in essence, is with replacement. What is the probability that the two pads selected are both improperly bound?

Solution

Let I denote improperly bound. The problem is to determine

$$P(I_1 \cap I_2) = ?$$

The probability of I = .03, or 3% are improperly bound. Because the sampling is done with replacement, the two events are independent. Hence,

$$P(I_1 \cap I_2) = P(I_1) \cdot P(I_2) = (.03)(.03) = .0009$$

TABLE 4.6

Contingency Table of Data from Independent Events

	D	*E*	
A	8	12	20
B	20	30	50
C	6	9	15
	34	51	85

Most probability matrices contain variables that are not independent. If a probability matrix contains independent events, the special law of multiplication can be applied. If not, the special law cannot be used. In Section 4.7 we explore a technique for determining whether events are independent. Table 4.6 contains data from independent events.

DEMONSTRATION PROBLEM 4.8

Use the data from Table 4.6 and the special law of multiplication to find $P(B \cap D)$.

Solution

$$P(B \cap D) = P(B) \cdot P(D) = \frac{50}{85} \cdot \frac{34}{85} = .2353$$

This approach works *only* for contingency tables and probability matrices in which the variable along one side of the matrix is *independent* of the variable along the other side of the matrix. Note that the answer obtained by using the formula is the same as the answer obtained by using the cell information from Table 4.6.

$$P(B \cap D) = \frac{20}{85} = .2353$$

4.6 PROBLEMS

4.15 Use the values in the contingency table to solve the equations given.

	C	*D*	*E*	*F*
A	5	11	16	8
B	2	3	5	7

a. $P(A \cap E) =$ ______

b. $P(D \cap B) =$ ______

c. $P(D \cap E) =$ ______

d. $P(A \cap B) =$ ______

4.16 Use the values in the probability matrix to solve the equations given.

	D	*E*	*F*
A	.12	.13	.08
B	.18	.09	.04
C	.06	.24	.06

a. $P(E \cap B) =$ ______
b. $P(C \cap F) =$ ______
c. $P(E \cap D) =$ ______

4.17 **a.** A batch of 50 parts contains six defects. If two parts are drawn randomly one at a time without replacement, what is the probability that both parts are defective?
b. If this experiment is repeated, with replacement, what is the probability that both parts are defective?

4.18 According to the nonprofit group Zero Population Growth, 78% of the U.S. population now lives in urban areas. Scientists at Princeton University and the University of Wisconsin report that about 15% of all American adults care for ill relatives. Suppose that 11% of adults living in urban areas care for ill relatives.
a. Use the general law of multiplication to determine the probability of randomly selecting an adult from the U.S. population who lives in an urban area and is caring for an ill relative.
b. What is the probability of randomly selecting an adult from the U.S. population who lives in an urban area and does not care for an ill relative?
c. Construct a probability matrix and show where the answer to this problem lies in the matrix.
d. From the probability matrix, determine the probability that an adult lives in a nonurban area and cares for an ill relative.

4.19 A study by Peter D. Hart Research Associates for the Nasdaq Stock Market revealed that 43% of all American adults are stockholders. In addition, the study determined that 75% of all American adult stockholders have some college education. Suppose 37% of all American adults have some college education. An American adult is randomly selected.
a. What is the probability that the adult does not own stock?
b. What is the probability that the adult owns stock and has some college education?
c. What is the probability that the adult owns stock or has some college education?
d. What is the probability that the adult has neither some college education nor owns stock?
e. What is the probability that the adult does not own stock or has no college education?
f. What is the probability that the adult has some college education and owns no stock?

4.20 According to the Consumer Electronics Manufacturers Association, 10% of all U.S. households have a fax machine and 52% have a personal computer. Suppose 91% of all U.S. households having a fax machine have a personal computer. A U.S. household is randomly selected.
a. What is the probability that the household has a fax machine and a personal computer?
b. What is the probability that the household has a fax machine or a personal computer?
c. What is the probability that the household has a fax machine and does not have a personal computer?
d. What is the probability that the household has neither a fax machine nor a personal computer?
e. What is the probability that the household does not have a fax machine and does have a personal computer?

4.21 A study by Becker Associates, a San Diego travel consultant, found that 30% of the traveling public said that their flight selections are influenced by perceptions of airline safety. Thirty-nine percent of the traveling public wants to know the age of the aircraft. Suppose 87% of the traveling public who say that their flight selections are influenced by perceptions of airline safety wants to know the age of the aircraft.

a. What is the probability of randomly selecting a member of the traveling public and finding out that she says that flight selection is influenced by perceptions of airline safety and she does not want to know the age of the aircraft?

b. What is the probability of randomly selecting a member of the traveling public and finding out that she says that flight selection is neither influenced by perceptions of airline safety nor does she want to know the age of the aircraft?

c. What is the probability of randomly selecting a member of the traveling public and finding out that he says that flight selection is not influenced by perceptions of airline safety and he wants to know the age of the aircraft?

4.22 The U.S. Energy Department states that 60% of all U.S. households have ceiling fans. In addition, 29% of all U.S. households have an outdoor grill. Suppose 13% of all U.S. households have both a ceiling fan and an outdoor grill. A U.S. household is randomly selected.

a. What is the probability that the household has a ceiling fan or an outdoor grill?

b. What is the probability that the household has neither a ceiling fan nor an outdoor grill?

c. What is the probability that the household does not have a ceiling fan and does have an outdoor grill?

d. What is the probability that the household does have a ceiling fan and does not have an outdoor grill?

4.7 CONDITIONAL PROBABILITY

Video

Conditional probabilities are computed based on the prior knowledge that a business researcher has on one of the two events being studied. If X, Y are two events, the conditional probability of X occurring given that Y is known or has occurred is expressed as $P(X|Y)$ and is given in the *law of conditional probability.*

LAW OF CONDITIONAL PROBABILITY

$$P(X|Y) = \frac{P(X \cap Y)}{P(Y)} = \frac{P(X) \cdot P(Y|X)}{P(Y)}$$

The conditional probability of $(X|Y)$ is the probability that X will occur given Y. The formula for conditional probability is derived by dividing both sides of the general law of multiplication by $P(Y)$.

In the study by Yankelovich Partners to determine what changes in office design would improve productivity, 70% of the respondents believed noise reduction would improve productivity and 67% said increased storage space would improve productivity. In addition, suppose 56% of the respondents believed both noise reduction and increased storage space would improve productivity. A worker is selected randomly and asked about changes in office design. This worker believes that noise reduction would improve productivity. What is the probability that this worker believes increased storage space would improve productivity? That is, what is the probability that a randomly selected person believes storage space would improve productivity *given that* he or she believes noise reduction improves productivity? In symbols, the question is

$$P(\text{S}|\text{N}) = ?$$

Note that the given part of the information is listed to the right of the vertical line in the conditional probability. The formula solution is

$$P(\mathrm{S}|\mathrm{N}) = \frac{P(\mathrm{S} \cap \mathrm{N})}{P(\mathrm{N})}$$

but

$$P(\mathrm{N}) = .70 \quad \text{and} \quad P(\mathrm{S} \cap \mathrm{N}) = .56$$

therefore

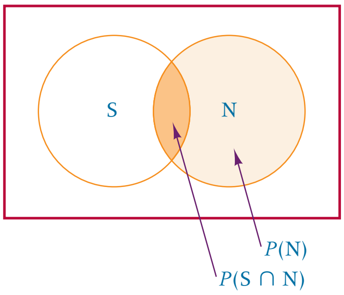
$$P(\mathrm{S}|\mathrm{N}) = \frac{P(\mathrm{S} \cap \mathrm{N})}{P(\mathrm{N})} = \frac{.56}{.70} = .80$$

Eighty percent of workers who believe noise reduction would improve productivity believe increased storage space would improve productivity.

Note in Figure 4.11 that the area for N in the Venn diagram is completely shaded because it is given that the worker believes noise reduction will improve productivity. Also notice that the intersection of N and S is more heavily shaded. This portion of noise reduction includes increased storage space. It is the only part of increased storage space that is in noise reduction, and because the person is known to favor noise reduction, it is the only area of interest that includes increased storage space.

FIGURE 4.11

Conditional Probability of Increased Storage Space Given Noise Reduction



Examine the probability matrix in Table 4.7 for the office design problem. None of the probabilities given in the matrix are conditional probabilities. To reiterate what has been previously stated, a probability matrix contains only two types of probabilities, marginal and joint. The cell values are all joint probabilities and the subtotals in the margins are marginal probabilities. How are conditional probabilities determined from a probability matrix? The law of conditional probabilities shows that a conditional probability is computed by dividing the joint probability by the marginal probability. Thus, the probability matrix has all the necessary information to solve for a conditional probability.

What is the probability that a randomly selected worker believes noise reduction would not improve productivity given that the worker does believe increased storage space would improve productivity? That is,

$$P(\text{not N}|\mathrm{S}) = ?$$

The law of conditional probability states that

$$P(\text{not N}|\mathrm{S}) = \frac{P(\text{not N} \cap \mathrm{S})}{P(\mathrm{S})}$$

Notice that because S is given, we are interested only in the column that is shaded in Table 4.7, which is the Yes column for increased storage space. The marginal probability, $P(\mathrm{S})$, is the total of this column and is found in the margin at the bottom of the table as .67. $P(\text{not N} \cap \mathrm{S})$ is found as the intersection of No for noise and Yes for storage. This value is .11. Hence, $P(\text{not N} \cap \mathrm{S})$ is .11. Therefore,

$$P(\text{not N}|\mathrm{S}) = \frac{P(\text{not N} \cap \mathrm{S})}{P(\mathrm{S})} = \frac{.11}{.67} = .164$$

TABLE 4.7

Office Design Problem Probability Matrix

		Increase Storage Space		
		Yes	*No*	
Noise Reduction	*Yes*	.56	.14	.70
	No	.11	.19	.30
		.67	.33	1.00

The second version of the conditional probability law formula is

$$P(X|Y) = \frac{P(X) \cdot P(Y \cdot X)}{P(Y)}$$

This version is more complex than the first version, $P(X \cap Y)/P(Y)$. However, sometimes the second version must be used because of the information given in the problem—for example, when solving for $P(X|Y)$ but $P(Y|X)$ is given. The second version of the formula is obtained from the first version by substituting the formula for $P(X \cap Y) = P(X) \cdot P(Y|X)$ into the first version.

As an example, in Section 4.6, data relating to women in the U.S. labor force were presented. Included in this information was the fact that 46% of the U.S. labor force is female and that 25% of the females in the U.S. labor force work part-time. In addition, 17.4% of all American laborers are known to be part-time workers. What is the probability that a randomly selected American worker is a woman if that person is known to be a part-time worker? Let W denote the event of selecting a woman and T denote the event of selecting a part-time worker. In symbols, the question to be answered is

$$P(\text{W}|\text{T}) = ?$$

The first form of the law of conditional probabilities is

$$P(\text{W}|\text{T}) = \frac{P(\text{W} \cap \text{T})}{P(\text{T})}$$

Note that this version of the law of conditional probabilities requires knowledge of the joint probability, $P(\text{W} \cap \text{T})$, which is not given here. We therefore try the second version of the law of conditional probabilities, which is

$$P(\text{W}|\text{T}) = \frac{P(\text{W}) \cdot P(\text{T}|\text{W})}{P(\text{T})}$$

For this version of the formula, everything is given in the problem.

$$P(\text{W}) = .46$$
$$P(\text{T}) = .174$$
$$P(\text{T}|\text{W}) = .25$$

The probability of a laborer being a woman given that the person works part-time can now be computed.

$$P(\text{W}|\text{T}) = \frac{P(\text{W}) \cdot P(\text{T}|\text{W})}{P(\text{T})} = \frac{(.46)(.25)}{(.174)} = .661$$

Hence, 66.1% of the part-time workers are women.

In general, this second version of the law of conditional probabilities is likely to be used for solving $P(X|Y)$ when $P(X \cap Y)$ is unknown but $P(Y|X)$ is known.

DEMONSTRATION PROBLEM 4.9

The data from the executive interviews given in Demonstration Problem 4.2 are repeated here. Use these data to find:

a. $P(\text{B}|\text{F})$

b. $P(\text{G}|\text{C})$

c. $P(\text{D}|\text{F})$

PROBABILITY MATRIX

		Geographic Location				
		Northeast D	Southeast E	Midwest F	West G	
Industry Type	Finance A	.12	.05	.04	.07	.28
	Manufacturing B	.15	.03	.11	.06	.35
	Communications C	.14	.09	.06	.08	.37
		.41	.17	.21	.21	1.00

Solution

a. $$P(\text{B}|\text{F}) = \frac{P(\text{B} \cap \text{F})}{P(\text{F})} = \frac{.11}{.21} = .524$$

Determining conditional probabilities from a probability matrix by using the formula is a relatively painless process. In this case, the joint probability, $P(\text{B} \cap \text{F})$, appears in a cell of the matrix (.11); the marginal probability, $P(\text{F})$, appears in a margin (.21). Bringing these two probabilities together by formula produces the answer, $.11/.21 = .524$. This answer means that 52.4% of the Midwest executives (the F values) are in manufacturing (the B values).

b. $$P(\text{G}|\text{C}) = \frac{P(\text{G} \cap \text{C})}{P(\text{C})} = \frac{.08}{.37} = .216$$

This result means that 21.6% of the responding communications industry executives, (C) are from the West (G).

c. $$P(\text{D}|\text{F}) = \frac{P(\text{D} \cap \text{F})}{P(\text{F})} = \frac{.00}{.21} = .00$$

Because D and F are mutually exclusive, $P(\text{D} \cap \text{F})$ is zero and so is $P(\text{D}|\text{F})$. The rationale behind $P(\text{D}|\text{F}) = 0$ is that, if F is given (the respondent is known to be located in the Midwest), the respondent could not be located in D (the Northeast).

Independent Events

INDEPENDENT EVENTS *X*, *Y*

To test to determine if X and Y are independent events, the following must be true.

$$P(X|Y) = P(X) \quad \text{and} \quad P(Y|X) = P(Y)$$

In each equation, it does not matter that X or Y is given because X and Y are *independent*. When X and Y are independent, the conditional probability is solved as a marginal probability.

Sometimes, it is important to test a contingency table of raw data to determine whether events are independent. If *any* combination of two events from the different sides of the matrix fail the test, $P(X|Y) = P(X)$, the matrix does not contain independent events.

DEMONSTRATION PROBLEM 4.10

Test the matrix for the 200 executive responses to determine whether industry type is independent of geographic location.

STATISTICS IN BUSINESS TODAY

Newspaper Advertising Reading Habits of Canadians

A national survey by Ipsos Reid for the Canadian Newspaper Association reveals some interesting statistics about newspaper advertising reading habits of Canadians. Sixty-six percent of Canadians say that they enjoy reading the page advertising and the product inserts that come with a newspaper. The percentage is higher for women (70%) than men (62%), but 73% of households with children enjoy doing so. While the percentage of those over 55 years of age who enjoy reading such ads is 71%, the percentage is only 55% for those in the 18-to-34-year-old category. These percentages decrease with increases in education as revealed by the fact that while 70% of those with a high school education enjoy reading such ads, only 55% of those having a university degree do so. Canadians living in the Atlantic region lead the country in this regard with 74%, in contrast to those living in British Columbia (63%) and Quebec (62%).

These facts can be converted to probabilities: The probability that a Canadian enjoys reading such ads is .66. Many of the other statistics represent conditional probabilities. For example, the probability that a Canadian enjoys such ads given that the Canadian is a woman is .70; and the probability that a Canadian enjoys such ads given that the Canadian has a college degree is .55. About 13% of the Canadian population resides in British Columbia. From this and from the conditional probability that a Canadian enjoys such ads given that they live in British Columbia (.63), one can compute the joint probability that a randomly selected Canadian enjoys such ads and lives in British Columbia (.13)(.63) = .0819. That is, 8.19% of all Canadians live in British Columbia and enjoy such ads.

RAW VALUES MATRIX

		Geographic Location				
		Northeast D	Southeast E	Midwest F	West G	
Industry Type	Finance A	24	10	8	14	56
	Manufacturing B	30	6	22	12	70
	Communications C	28	18	12	16	74
		82	34	42	42	200

Solution

Select one industry and one geographic location (say, A—Finance and G—West). Does $P(A|G) = P(A)$?

$$P(A|G) = \frac{14}{42} \text{ and } P(A) = \frac{56}{200}$$

Does 14/42 = 56/200? No, .33 ≠ .28. Industry and geographic location are not independent because at least one exception to the test is present.

DEMONSTRATION PROBLEM 4.11

Determine whether the contingency table shown as Table 4.6 and repeated here contains independent events.

	D	E	
A	8	12	20
B	20	30	50
C	6	9	15
	34	51	85

Solution

Check the first cell in the matrix to find whether $P(A|D) = P(A)$.

$$P(A|D) = \frac{8}{34} = .2353$$

$$P(A) = \frac{20}{85} = .2353$$

The checking process must continue until all the events are determined to be independent. In this matrix, all the possibilities check out. Thus, Table 4.6 contains independent events.

Video

4.7 PROBLEMS

4.23 Use the values in the contingency table to solve the equations given.

	E	*F*	*G*
A	15	12	8
B	11	17	19
C	21	32	27
D	18	13	12

a. $P(G|A) =$ _____
b. $P(B|F) =$ _____
c. $P(C|E) =$ _____
d. $P(E|G) =$ _____

4.24 Use the values in the probability matrix to solve the equations given.

	C	*D*
A	.36	.44
B	.11	.09

a. $P(C|A) =$ _____
b. $P(B|D) =$ _____
c. $P(A|B) =$ _____

4.25 The results of a survey asking, "Do you have a calculator and/or a computer in your home?" follow.

		Calculator		
		Yes	*No*	
Computer	*Yes*	46	3	49
	No	11	15	26
		57	18	75

Is the variable "calculator" independent of the variable "computer"? Why or why not?

4.26 In a recent year, business failures in the United States numbered 83,384, according to Dun & Bradstreet. The construction industry accounted for 10,867 of these business failures. The South Atlantic states accounted for 8,010 of the business failures. Suppose that 1,258 of all business failures were construction businesses located in the South Atlantic states. A failed business is randomly selected from this list of business failures.

a. What is the probability that the business is located in the South Atlantic states?

b. What is the probability that the business is in the construction industry or located in the South Atlantic states?

c. What is the probability that the business is in the construction industry if it is known that the business is located in the South Atlantic states?

d. What is the probability that the business is located in the South Atlantic states if it is known that the business is a construction business?

e. What is the probability that the business is not located in the South Atlantic states if it is known that the business is not a construction business?

f. Given that the business is a construction business, what is the probability that the business is not located in the South Atlantic states?

4.27 Arthur Andersen Enterprise Group/National Small Business United, Washington, conducted a national survey of small-business owners to determine the challenges for growth for their businesses. The top challenge, selected by 46% of the small-business owners, was the economy. A close second was finding qualified workers (37%). Suppose 15% of the small-business owners selected both the economy and finding qualified workers as challenges for growth. A small-business owner is randomly selected.

a. What is the probability that the owner believes the economy is a challenge for growth if the owner believes that finding qualified workers is a challenge for growth?

b. What is the probability that the owner believes that finding qualified workers is a challenge for growth if the owner believes that the economy is a challenge for growth?

c. Given that the owner does not select the economy as a challenge for growth, what is the probability that the owner believes that finding qualified workers is a challenge for growth?

d. What is the probability that the owner believes neither that the economy is a challenge for growth nor that finding qualified workers is a challenge for growth?

4.28 According to a survey published by ComPsych Corporation, 54% of all workers read e-mail while they are talking on the phone. Suppose that 20% of those who read e-mail while they are talking on the phone write personal "to-do" lists during meetings. Assuming that these figures are true for all workers, if a worker is randomly selected, determine the following probabilities:

a. The worker reads e-mail while talking on the phone and writes personal "to-do" lists during meetings.

b. The worker does not write personal "to-do" lists given that he reads e-mail while talking on the phone.

c. The worker does not write personal "to-do" lists and does read e-mail while talking on the phone.

4.29 *Accounting Today* reported that 37% of accountants purchase their computer hardware by mail order direct and that 54% purchase their computer software by mail order direct. Suppose that 97% of the accountants who purchase their computer hardware by mail order direct purchase their computer software by mail order direct. If an accountant is randomly selected, determine the following probabilities:

a. The accountant does not purchase his computer software by mail order direct given that he does purchase his computer hardware by mail order direct.

b. The accountant does purchase his computer software by mail order direct given that he does not purchase his computer hardware by mail order direct.

c. The accountant does not purchase his computer hardware by mail order direct if it is known that he does purchase his computer software by mail order direct.

d. The accountant does not purchase his computer hardware by mail order direct if it is known that he does not purchase his computer software by mail order direct.

4.30 In a study undertaken by Catalyst, 43% of women senior executives agreed or strongly agreed that a lack of role models was a barrier to their career development. In addition, 46% agreed or strongly agreed that gender-based stereotypes were barriers to their career advancement. Suppose 77% of those who agreed or strongly agreed that gender-based stereotypes were barriers to their career advancement agreed or strongly agreed that the lack of role models was a barrier to their career development. If one of these female senior executives is randomly selected, determine the following probabilities:

a. What is the probability that the senior executive does not agree or strongly agree that a lack of role models was a barrier to her career development given that she does agree or strongly agree that gender-based stereotypes were barriers to her career development?

b. What is the probability that the senior executive does not agree or strongly agree that gender-based stereotypes were barriers to her career development given that she does agree or strongly agree that the lack of role models was a barrier to her career development?

c. If it is known that the senior executive does not agree or strongly agree that gender-based stereotypes were barriers to her career development, what is the probability that she does not agree or strongly agree that the lack of role models was a barrier to her career development?

4.8 REVISION OF PROBABILITIES: BAYES' RULE

An extension to the conditional law of probabilities is Bayes' rule, which was developed by and named for Thomas Bayes (1702–1761). **Bayes' rule** is *a formula that extends the use of the law of conditional probabilities to allow revision of original probabilities with new information.*

BAYES' RULE

$$P(X_i|Y) = \frac{P(X_i) \cdot P(Y|X_i)}{P(X_1) \cdot P(Y|X_1) + P(X_2) \cdot P(Y|X_2) + \cdots + P(X_n) \cdot P(Y|X_n)}$$

Recall that the law of conditional probability for

$$P(X_i|Y)$$

is

$$P(X_i|Y) = \frac{P(X_i) \cdot P(Y|X_i)}{P(Y)}$$

Compare Bayes' rule to this law of conditional probability. The numerators of Bayes' rule and the law of conditional probability are the same—the intersection of X_i and Y shown in the form of the general rule of multiplication. The new feature that Bayes' rule uses is found in the denominator of the rule:

$$P(X_1) \cdot P(Y|X_1) + P(X_2) \cdot P(Y|X_2) + \cdots + P(X_n) \cdot P(Y|X_n)$$

The denominator of Bayes' rule includes a product expression (intersection) for every partition in the sample space, Y, including the event (X_i) itself. The denominator is thus a collective exhaustive listing of mutually exclusive outcomes of Y. This denominator is sometimes referred to as the "total probability formula." It represents a weighted average of the conditional probabilities, with the weights being the prior probabilities of the corresponding event.

TABLE 4.8

Bayesian Table for Revision of Ribbon Problem Probabilities

Event	Prior Probability $P(E_i)$	Conditional Probability $P(d\|E_i)$	Joint Probability $P(E_i \cap d)$	Posterior or Revised Probability
Alamo	.65	.08	.052	$\frac{.052}{.094} = .553$
South Jersey	.35	.12	.042	$\frac{.042}{.094} = .447$
			$P(\text{defective}) = .094$	

By expressing the law of conditional probabilities in this new way, Bayes' rule enables the statistician to make new and different applications using conditional probabilities. In particular, statisticians use Bayes' rule to "revise" probabilities in light of new information.

A particular type of printer ribbon is produced by only two companies, Alamo Ribbon Company and South Jersey Products. Suppose Alamo produces 65% of the ribbons and that South Jersey produces 35%. Eight percent of the ribbons produced by Alamo are defective and 12% of the South Jersey ribbons are defective. A customer purchases a new ribbon. What is the probability that Alamo produced the ribbon? What is the probability that South Jersey produced the ribbon? The ribbon is tested, and it is defective. Now what is the probability that Alamo produced the ribbon? That South Jersey produced the ribbon?

The probability was .65 that the ribbon came from Alamo and .35 that it came from South Jersey. These are called *prior* probabilities because they are based on the original information.

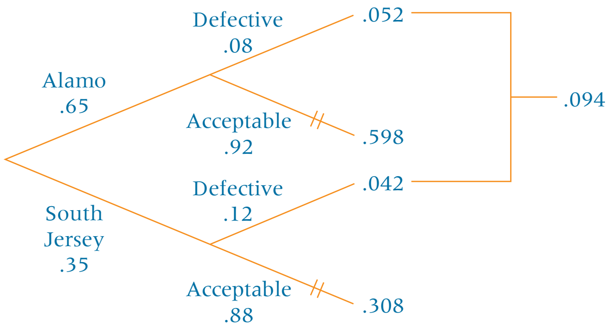
The new information that the ribbon is defective changes the probabilities because one company produces a higher percentage of defective ribbons than the other company does. How can this information be used to update or revise the original probabilities? Bayes' rule allows such updating. One way to lay out a revision of probabilities problem is to use a table. Table 4.8 shows the analysis for the ribbon problem.

The process begins with the prior probabilities: .65 Alamo and .35 South Jersey. These prior probabilities appear in the second column of Table 4.8. Because the product is found to be defective, the conditional probabilities, $P(\text{defective}|\text{Alamo})$ and $P(\text{defective}|\text{South Jersey})$ should be used. Eight percent of Alamo's ribbons are defective: $P(\text{defective}|\text{Alamo}) = .08$. Twelve percent of South Jersey's ribbons are defective: $P(\text{defective}|\text{South Jersey}) = .12$. These two conditional probabilities appear in the third column. Eight percent of Alamo's 65% of the ribbons are defective: $(.08)(.65) = .052$, or 5.2% of the total. This figure appears in the fourth column of Table 4.8; it is the joint probability of getting a ribbon that was made by Alamo and is defective. Because the purchased ribbon is defective, these are the only Alamo ribbons of interest. Twelve percent of South Jersey's 35% of the ribbons are defective. Multiplying these two percentages yields the joint probability of getting a South Jersey ribbon that is defective. This figure also appears in the fourth column of Table 4.8: $(.12)(.35) = .042$; that is, 4.2% of all ribbons are made by South Jersey and are defective. This percentage includes the only South Jersey ribbons of interest because the ribbon purchased is defective.

Column 4 is totaled to get .094, indicating that 9.4% of all ribbons are defective (Alamo and defective $= .052 +$ South Jersey and defective $= .042$). The other 90.6% of the ribbons, which are acceptable, are not of interest because the ribbon purchased is defective. To compute the fifth column, the posterior or revised probabilities, involves dividing each value in column 4 by the total of column 4. For Alamo, .052 of the total ribbons are Alamo and defective out of the total of .094 that are defective. Dividing .052 by .094 yields .553 as a revised probability that the purchased ribbon was made by Alamo. This probability is lower than the prior or original probability of .65 because fewer of Alamo's ribbons (as a percentage) are defective than those produced by South Jersey. The defective ribbon is now less likely to have come from Alamo than before the knowledge of the defective ribbon. South Jersey's probability is revised by dividing the .042 joint probability of the ribbon being made by South Jersey and defective by the total probability of the ribbon being defective (.094).

FIGURE 4.12

Tree Diagram for Ribbon Problem Probabilities



The result is .042/.094 = .447. The probability that the defective ribbon is from South Jersey increased because a higher percentage of South Jersey ribbons are defective.

Tree diagrams are another common way to solve Bayes' rule problems. Figure 4.12 shows the solution for the ribbon problem. Note that the tree diagram contains all possibilities, including both defective and acceptable ribbons. When new information is given, only the pertinent branches are selected and used. The joint probability values at the end of the appropriate branches are used to revise and compute the posterior possibilities. Using the total number of defective ribbons, .052 + .042 = .094, the calculation is as follows.

$$\text{Revised Probability: Alamo} = \frac{.052}{.094} = .553$$

$$\text{Revised Probability: South Jersey} = \frac{.042}{.094} = .447$$

DEMONSTRATION PROBLEM 4.12

Machines A, B, and C all produce the same two parts, X and Y. Of all the parts produced, machine A produces 60%, machine B produces 30%, and machine C produces 10%. In addition,

40% of the parts made by machine A are part X.

50% of the parts made by machine B are part X.

70% of the parts made by machine C are part X.

A part produced by this company is randomly sampled and is determined to be an X part. With the knowledge that it is an X part, revise the probabilities that the part came from machine A, B, or C.

Solution

The prior probability of the part coming from machine A is .60, because machine A produces 60% of all parts. The prior probability is .30 that the part came from B and .10 that it came from C. These prior probabilities are more pertinent if nothing is known about the part. However, the part is known to be an X part. The conditional probabilities show that different machines produce different proportions of X parts. For example, .40 of the parts made by machine A are X parts, but .50 of the parts made by machine B and .70 of the parts made by machine C are X parts. It makes sense that the probability of the part coming from machine C would increase and that the probability that the part was made on machine A would decrease because the part is an X part.

The following table shows how the prior probabilities, conditional probabilities, joint probabilities, and marginal probability, $P(X)$, can be used to revise the prior probabilities to obtain posterior probabilities.

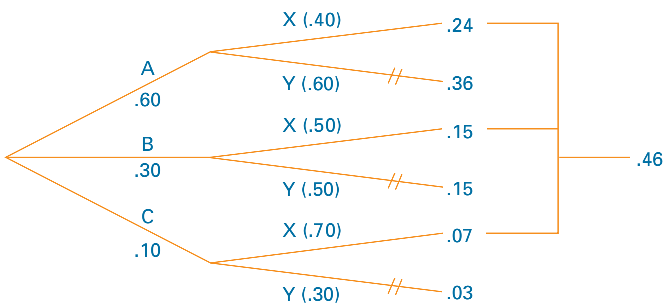
Event	Prior $P(E_i)$	Conditional $P(X \mid E_i)$	Joint $P(X \cap E_i)$	Posterior
A	.60	.40	$(.60)(.40) = .24$	$\frac{.24}{.46} = .52$
B	.30	.50	.15	$\frac{.15}{.46} = .33$
C	.10	.70	.07	$\frac{.07}{.46} = .15$
			$P(X) = .46$	

After the probabilities are revised, it is apparent that the probability of the part being made at machine A decreased and that the probabilities that the part was made at machines B and C increased. A tree diagram presents another view of this problem.

Revised Probabilities: Machine A: $\frac{.24}{.46} = .52$

Machine B: $\frac{.15}{.46} = .33$

Machine C: $\frac{.07}{.46} = .15$



4.8 PROBLEMS

4.31 In a manufacturing plant, machine A produces 10% of a certain product, machine B produces 40% of this product, and machine C produces 50% of this product. Five percent of machine A products are defective, 12% of machine B products are defective, and 8% of machine C products are defective. The company inspector has just sampled a product from this plant and has found it to be defective. Determine the revised probabilities that the sampled product was produced by machine A, machine B, or machine C.

4.32 Alex, Alicia, and Juan fill orders in a fast-food restaurant. Alex incorrectly fills 20% of the orders he takes. Alicia incorrectly fills 12% of the orders she takes. Juan incorrectly fills 5% of the orders he takes. Alex fills 30% of all orders, Alicia fills 45% of all orders, and Juan fills 25% of all orders. An order has just been filled.

a. What is the probability that Alicia filled the order?

b. If the order was filled by Juan, what is the probability that it was filled correctly?

c. Who filled the order is unknown, but the order was filled incorrectly. What are the revised probabilities that Alex, Alicia, or Juan filled the order?

d. Who filled the order is unknown, but the order was filled correctly. What are the revised probabilities that Alex, Alicia, or Juan filled the order?

4.33 In a small town, two lawn companies fertilize lawns during the summer. Tri-State Lawn Service has 72% of the market. Thirty percent of the lawns fertilized by Tri-State could be rated as very healthy one month after service. Greenchem has the other 28% of the market. Twenty percent of the lawns fertilized by Greenchem could be rated as very healthy one month after service. A lawn that has been treated with fertilizer by one of these companies within the last month is selected randomly. If the lawn is rated as very healthy, what are the revised probabilities that Tri-State or Greenchem treated the lawn?

4.34 Suppose 70% of all companies are classified as small companies and the rest as large companies. Suppose further, 82% of large companies provide training to employees, but only 18% of small companies provide training. A company is randomly selected without knowing if it is a large or small company; however, it is determined that the company provides training to employees. What are the prior probabilities that the company is a large company or a small company? What are the revised probabilities that the company is large or small? Based on your analysis, what is the overall percentage of companies that offer training?

Decision Dilemma Solved

Equity of the Sexes in the Workplace

The client company data given in the Decision Dilemma are displayed in a raw values matrix form. Using the techniques presented in this chapter, it is possible to statistically answer the managerial questions. If a worker is randomly selected from the 155 employees, the probability that the worker is a woman, $P(W)$, is 55/155, or .355. This marginal probability indicates that roughly 35.5% of all employees of the client company are women. Given that the employee has a managerial position, the probability that the employee is a woman, $P(W|M)$ is 3/11, or .273. The proportion of managers at the company who are women is lower than the proportion of all workers at the company who are women. Several factors might be related to this discrepancy, some of which may be defensible by the company—including experience, education, and prior history of success—and some may not.

Suppose a technical employee is randomly selected for a bonus. What is the probability that a female would be selected given that the worker is a technical employee? That is, $P(F|T) = ?$ Applying the law of conditional probabilities to the raw values matrix given in the Decision Dilemma, $P(F|T) = 17/69 = .246$. Using the concept of complementary events, the probability that a man is selected given that the employee is a technical person is $1 - .246 = .754$. It is more than three times as likely that a randomly selected technical person is a male. If a woman were the one chosen for the bonus, a man could argue discrimination based on the mere probabilities. However, the company decision makers could then present documentation of the choice criteria based on productivity, technical suggestions, quality measures, and others.

Suppose a client company employee is randomly chosen to win a trip to Hawaii. The marginal probability that the winner is a professional is $P(P) = 44/155 = .284$. The probability that the winner is either a male or is a clerical worker, a union probability, is:

$$P(M \cup C) = P(M) + P(C) - P(M \cap C)$$

$$= \frac{100}{155} + \frac{31}{155} - \frac{9}{155} = \frac{122}{155} = .787$$

The probability of a male or clerical employee at the client company winning the trip is .787. The probability that the winner is a woman *and* a manager, a joint probability, is

$$P(F \cap M) = 3/155 = .019$$

There is less than a 2% chance that a female manager will be selected randomly as the trip winner.

What is the probability that the winner is from the technical group if it is known that the employee is a male? This conditional probability is as follows:

$$P(T|M) = 52/100 = .52.$$

Many other questions can be answered about the client company's human resource situation using probabilities.

The probability approach to a human resource pool is a factual, numerical approach to people selection taken without regard to individual talents, skills, and worth to the company. Of course, in most instances, many other considerations go into the hiring, promoting, and rewarding of workers besides the random draw of their name. However, company management should be aware that attacks on hiring, promotion, and reward practices are sometimes made using statistical analyses such as those presented here. It is not being argued here that management should base decisions merely on the probabilities within particular categories. Nevertheless, being aware of the probabilities, management can proceed to undergird their decisions with documented evidence of worker productivity and worth to the organization.

ETHICAL CONSIDERATIONS

One of the potential misuses of probability occurs when subjective probabilities are used. Most subjective probabilities are based on a person's feelings, intuition, or experience. Almost everyone has an opinion on something and is willing to share it. Although such probabilities are not strictly unethical to report, they can be misleading and disastrous to other decision makers. In addition, subjective probabilities leave the door open for unscrupulous people to overemphasize their point of view by manipulating the probability.

The decision maker should remember that the laws and rules of probability are for the "long run." If a coin is tossed, even though the probability of getting a head is .5, the result will be either a head or a tail. It isn't possible to get a half head. The probability of getting a head (.5) will probably work out in the long run, but in the short run an experiment might produce 10 tails in a row. Suppose the probability of striking oil on a geological formation is .10. This probability means that, in the long run, if the company drills enough holes on this type of formation, it should strike oil in about 10% of the holes. However, if the company has only enough money to drill one hole, it will either strike oil or have a dry hole. The probability figure of .10 may mean something different to the company that can afford to drill only one hole than to the company that can drill many hundreds. Classical probabilities could be used unethically to lure a company or client into a potential short-run investment with the expectation of getting at least something in return, when in actuality the investor will either win or lose. The oil company that drills only one hole will not get 10% back from the hole. It will either win or lose on the hole. Thus, classical probabilities open the door for unsubstantiated expectations, particularly in the short run.

SUMMARY

The study of probability addresses ways of assigning probabilities, types of probabilities, and laws of probabilities. Probabilities support the notion of inferential statistics. Using sample data to estimate and test hypotheses about population parameters is done with uncertainty. If samples are taken at random, probabilities can be assigned to outcomes of the inferential process.

Three methods of assigning probabilities are (1) the classical method, (2) the relative frequency of occurrence method, and (3) subjective probabilities. The classical method can assign probabilities a priori, or before the experiment takes place. It relies on the laws and rules of probability. The relative frequency of occurrence method assigns probabilities based on historical data or empirically derived data. Subjective probabilities are based on the feelings, knowledge, and experience of the person determining the probability.

Certain special types of events necessitate amendments to some of the laws of probability: mutually exclusive events and independent events. Mutually exclusive events are events that cannot occur at the same time, so the probability of their intersection is zero. With independent events, the occurrence of one has no impact or influence on the occurrence of the other. Certain experiments, such as those involving coins or dice, naturally produce independent events. Other experiments produce independent events when the experiment is conducted with replacement. If events are independent, the joint probability is computed by multiplying the marginal probabilities, which is a special case of the law of multiplication.

Three techniques for counting the possibilities in an experiment are the *mn* counting rule, the N^n possibilities, and combinations. The *mn* counting rule is used to determine how many total possible ways an experiment can occur in a series of sequential operations. The N^n formula is applied when sampling is being done with replacement or events are independent. Combinations are used to determine the possibilities when sampling is being done without replacement.

Four types of probability are marginal probability, conditional probability, joint probability, and union probability. The general law of addition is used to compute the probability of a union. The general law of multiplication is used to compute joint probabilities. The conditional law is used to compute conditional probabilities.

Bayes' rule is a method that can be used to revise probabilities when new information becomes available; it is a variation of the conditional law. Bayes' rule takes prior probabilities of events occurring and adjusts or revises those probabilities on the basis of information about what subsequently occurs.

KEY TERMS

Flash Cards

a priori
Bayes' rule
classical method of assigning probabilities
collectively exhaustive events
combinations
complement of a union
complement
conditional probability
elementary events
event
experiment
independent events
intersection
joint probability
marginal probability
mn counting rule
mutually exclusive events
probability matrix
relative frequency of occurrence
sample space
set notation
subjective probability
union
union probability

FORMULAS

Counting rule

$$mn$$

Sampling with replacement

$$N^n$$

Sampling without replacement

$$_NC_n$$

Combination formula

$$_NC_n = \binom{N}{n} = \frac{N!}{n!(N-n)!}$$

General law of addition

$$P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)$$

Special law of addition

$$P(X \cup Y) = P(X) + P(Y)$$

General law of multiplication

$$P(X \cap Y) = P(X) \cdot P(Y|X) = P(Y) \cdot P(X|Y)$$

Special law of multiplication

$$P(X \cap Y) = P(X) \cdot P(Y)$$

Law of conditional probability

$$P(X|Y) = \frac{P(X \cap Y)}{P(Y)} = \frac{P(X) \cdot P(Y|X)}{P(Y)}$$

Bayes' rule

$$P(X_i|Y) = \frac{P(X_i) \cdot P(Y|X_i)}{P(X_1) \cdot P(Y|X_1) + P(X_2) \cdot P(Y|X_2) + \cdots + P(X_n) \cdot P(Y|X_n)}$$

SUPPLEMENTARY PROBLEMS

CALCULATING THE STATISTICS

4.35 Use the values in the contingency table to solve the equations given.

		D	*E*
Variable 2	*A*	10	20
	B	15	5
	C	30	15

a. $P(E)$ = _____
b. $P(B \cup D)$ = _____
c. $P(A \cap E)$ = _____
d. $P(B|E)$ = _____
e. $P(A \cup B)$ = _____
f. $P(B \cap C)$ = _____
g. $P(D|C)$ = _____
h. $P(A|B)$ = _____
i. Are variables 1 and 2 independent? Why or why not?

4.36 Use the values in the contingency table to solve the equations given.

	D	*E*	*F*	*G*
A	3	9	7	12
B	8	4	6	4
C	10	5	3	7

a. $P(F \cap A)$ = _____
b. $P(A|B)$ =
c. $P(B)$ = _____
d. $P(E \cap F)$ = _____
e. $P(D|B)$ = _____
f. $P(B|D)$ = _____
g. $P(D \cup C)$ = _____
h. $P(F)$ = _____

4.37 The following probability matrix contains a breakdown on the age and gender of U.S. physicians in a recent year, as reported by the American Medical Association.

U.S. PHYSICIANS IN A RECENT YEAR

		Age (years)					
		<35	35–44	45–54	55–64	>65	
Gender	*Male*	.11	.20	.19	.12	.16	.78
	Female	.07	.08	.04	.02	.01	.22
		.18	.28	.23	.14	.17	1.00

a. What is the probability that one randomly selected physician is 35–44 years old?
b. What is the probability that one randomly selected physician is both a woman and 45–54 years old?
c. What is the probability that one randomly selected physician is a man or is 35–44 years old?
d. What is the probability that one randomly selected physician is less than 35 years old or 55–64 years old?
e. What is the probability that one randomly selected physician is a woman if she is 45–54 years old?
f. What is the probability that a randomly selected physician is neither a woman nor 55–64 years old?

TESTING YOUR UNDERSTANDING

4.38 Purchasing Survey asked purchasing professionals what sales traits impressed them most in a sales representative. Seventy-eight percent selected "thoroughness." Forty percent responded "knowledge of your own product." The purchasing professionals were allowed to list more than one trait. Suppose 27% of the purchasing professionals listed both "thoroughness" and "knowledge of your own product" as sales traits that impressed them most. A purchasing professional is randomly sampled.

- **a.** What is the probability that the professional selected "thoroughness" or "knowledge of your own product"?
- **b.** What is the probability that the professional selected neither "thoroughness" nor "knowledge of your own product"?
- **c.** If it is known that the professional selected "thoroughness," what is the probability that the professional selected "knowledge of your own product"?
- **d.** What is the probability that the professional did not select "thoroughness" and did select "knowledge of your own product"?

4.39 The U.S. Bureau of Labor Statistics publishes data on the benefits offered by small companies to their employees. Only 42% offer retirement plans while 61% offer life insurance. Suppose 33% offer both retirement plans and life insurance as benefits. If a small company is randomly selected, determine the following probabilties:

- **a.** The company offers a retirement plan given that they offer life insurance.
- **b.** The company offers life insurance given that they offer a retirement plan.
- **c.** The company offers life insurance or a retirement plan.
- **d.** The company offers a retirement plan and does not offer life insurance.
- **e.** The company does not offer life insurance if it is known that they offer a retirement plan.

4.40 According to Link Resources, 16% of the U.S. population is technology driven. However, these figures vary by region. For example, in the West the figure is 20% and in the Northeast the figure is 17%. Twenty-one percent of the U.S. population in general is in the West and 20% of the U.S. population is in the Northeast. Suppose an American is chosen randomly.

- **a.** What is the probability that the person lives in the West and is a technology-driven person?
- **b.** What is the probability that the person lives in the Northeast and is a technology-driven person?
- **c.** Suppose the chosen person is known to be technology-driven. What is the probability that the person lives in the West?
- **d.** Suppose the chosen person is known not to be technology-driven. What is the probability that the person lives in the Northeast?
- **e.** Suppose the chosen person is known to be technology-driven. What is the probability that the person lives in neither the West nor the Northeast?

4.41 In a certain city, 30% of the families have a MasterCard, 20% have an American Express card, and 25% have a Visa card. Eight percent of the families have both a MasterCard and an American Express card. Twelve percent have both a Visa card and a MasterCard. Six percent have both an American Express card and a Visa card.

- **a.** What is the probability of selecting a family that has either a Visa card or an American Express card?
- **b.** If a family has a MasterCard, what is the probability that it has a Visa card?
- **c.** If a family has a Visa card, what is the probability that it has a MasterCard?
- **d.** Is possession of a Visa card independent of possession of a MasterCard? Why or why not?
- **e.** Is possession of an American Express card mutually exclusive of possession of a Visa card?

4.42 A few years ago, a survey commissioned by *The World Almanac* and *Maturity News* Service reported that 51% of the respondents did not believe the Social Security system will be secure in 20 years. Of the respondents who were age 45 or older, 70% believed the system will be secure in 20 years. Of the people surveyed, 57% were under age 45. One respondent is selected randomly.

- **a.** What is the probability that the person is age 45 or older?
- **b.** What is the probability that the person is younger than age 45 and believes that the Social Security system will be secure in 20 years?
- **c.** If the person selected believes the Social Security system will be secure in 20 years, what is the probability that the person is 45 years old or older?
- **d.** What is the probability that the person is younger than age 45 or believes the Social Security system will not be secure in 20 years?

4.43 A telephone survey conducted by the Maritz Marketing Research company found that 43% of Americans expect to save more money next year than they saved last year. Forty-five percent of those surveyed plan to reduce debt next year. Of those who expect to save more money next year, 81% plan to reduce debt next year. An American is selected randomly.

- **a.** What is the probability that this person expects to save more money next year and plans to reduce debt next year?
- **b.** What is the probability that this person expects to save more money next year or plans to reduce debt next year?
- **c.** What is the probability that this person neither expects to save more money next year nor plans to reduce debt next year?
- **d.** What is the probability that this person expects to save more money next year and does not plan to reduce debt next year?

4.44 The Steelcase Workplace Index studied the types of work-related activities that Americans did while on vacation in the summer. Among other things, 40% read work-related material. Thirty-four percent checked in with the boss.

Respondents to the study were allowed to select more than one activity. Suppose that of those who read work-related material, 78% checked in with the boss. One of these survey respondents is selected randomly.

a. What is the probability that while on vacation this respondent checked in with the boss and read work-related material?
b. What is the probability that while on vacation this respondent neither read work-related material nor checked in with the boss?
c. What is the probability that while on vacation this respondent read work-related material given that the respondent checked in with the boss?
d. What is the probability that while on vacation this respondent did not check in with the boss given that the respondent read work-related material?
e. What is the probability that while on vacation this respondent did not check in with the boss given that the respondent did not read work-related material?
f. Construct a probability matrix for this problem.

4.45 A study on ethics in the workplace by the Ethics Resource Center and Kronos, Inc., revealed that 35% of employees admit to keeping quiet when they see coworker misconduct. Suppose 75% of employees who admit to keeping quiet when they see coworker misconduct call in sick when they are well. In addition, suppose that 40% of the employees who call in sick when they are well admit to keeping quiet when they see coworker misconduct. If an employee is randomly selected, determine the following probabilities:

a. The employee calls in sick when well and admits to keeping quiet when seeing coworker misconduct.
b. The employee admits to keeping quiet when seeing coworker misconduct or calls in sick when well.
c. Given that the employee calls in sick when well, he or she does not keep quiet when seeing coworker misconduct.
d. The employee neither keeps quiet when seeing coworker misconduct nor calls in sick when well.
e. The employee admits to keeping quiet when seeing coworker misconduct and does not call in sick when well.

4.46 Health Rights Hotline published the results of a survey of 2,400 people in Northern California in which consumers were asked to share their complaints about managed care. The number one complaint was denial of care, with 17% of the participating consumers selecting it. Several other complaints were noted, including inappropriate care (14%), customer service (14%), payment disputes (11%), specialty care (10%), delays in getting care (8%), and prescription drugs (7%). These complaint categories are mutually exclusive. Assume that the results of this survey can be inferred to all managed care consumers. If a managed care consumer is randomly selected, determine the following probabilities:

a. The consumer complains about payment disputes or specialty care.
b. The consumer complains about prescription drugs and customer service.
c. The consumer complains about inappropriate care given that the consumer complains about specialty care.
d. The consumer does not complain about delays in getting care nor does the consumer complain about payment disputes.

4.47 Companies use employee training for various reasons, including employee loyalty, certification, quality, and process improvement. In a national survey of companies, BI Learning Systems reported that 56% of the responding companies named employee retention as a top reason for training. Suppose 36% of the companies replied that they use training for process improvement and for employee retention. In addition, suppose that of the companies that use training for process improvement, 90% use training for employee retention. A company that uses training is randomly selected.

a. What is the probability that the company uses training for employee retention and not for process improvement?
b. If it is known that the company uses training for employee retention, what is the probability that it uses training for process improvement?
c. What is the probability that the company uses training for process improvement?
d. What is the probability that the company uses training for employee retention or process improvement?
e. What is the probability that the company neither uses training for employee retention nor uses training for process improvement?
f. Suppose it is known that the company does not use training for process improvement. What is the probability that the company does use training for employee retention?

4.48 Pitney Bowes surveyed 302 directors and vice presidents of marketing at large and midsized U.S. companies to determine what they believe is the best vehicle for educating decision makers on complex issues in selling products and services. The highest percentage of companies chose direct mail/catalogs, followed by direct sales/sales rep. Direct mail/catalogs was selected by 38% of the companies. None of the companies selected both direct mail/catalogs and direct sales/sales rep. Suppose also that 41% selected neither direct mail/catalogs nor direct sales/sales rep. If one of these companies is selected randomly and their top marketing person interviewed about this matter, determine the following probabilities:

a. The marketing person selected direct mail/catalogs and did not select direct sales/sales rep.
b. The marketing person selected direct sales/sales rep.
c. The marketing person selected direct sales/sales rep given that the person selected direct mail/catalogs.
d. The marketing person did not select direct mail/catalogs given that the person did not select direct sales/sales rep.

4.49 In a study of incentives used by companies to retain mature workers by The Conference Board, it was reported that 41% use flexible work arrangements. Suppose that of those companies that do not use flexible work arrangements, 10% give time off for volunteerism. In addition, suppose that of those companies that use flexible work arrangements, 60% give time off for volunteerism. If a company is randomly selected, determine the following probabilities:

a. The company uses flexible work arrangements or gives time off for volunteerism.

b. The company uses flexible work arrangements and does not give time off for volunteerism.

c. Given that the company does not give time off for volunteerism, the company uses flexible work arrangements.

d. The company does not use flexible work arrangements given that the company does give time off for volunteerism.

e. The company does not use flexible work arrangements or the company does not give time off for volunteerism.

4.50 A small independent physicians' practice has three doctors. Dr. Sarabia sees 41% of the patients, Dr. Tran sees 32%, and Dr. Jackson sees the rest. Dr. Sarabia requests blood tests on 5% of her patients, Dr. Tran requests blood tests on 8% of his patients, and Dr. Jackson requests blood tests on 6% of her patients. An auditor randomly selects a patient from the past week and discovers that the patient had a blood test as a result of the physician visit. Knowing this information, what is the probability that the patient saw Dr. Sarabia? For what percentage of all patients at this practice are blood tests requested?

4.51 A survey by the Arthur Andersen Enterprise Group/National Small Business United attempted to determine what the leading challenges are for the growth and survival of small businesses. Although the economy and finding qualified workers were the leading challenges, several others were listed in the results of the study, including regulations, listed by 30% of the companies, and the tax burden, listed by 35%. Suppose that 71% of the companies listing regulations as a challenge listed the tax burden as a challenge. Assume these percentages hold for all small businesses. If a small business is randomly selected, determine the following probabilities:

a. The small business lists both the tax burden and regulations as a challenge.

b. The small business lists either the tax burden or regulations as a challenge.

c. The small business lists either the tax burden or regulations but not both as a challenge.

d. The small business lists regulations as a challenge given that it lists the tax burden as a challenge.

e. The small business does not list regulations as a challenge given that it lists the tax burden as a challenge.

f. The small business does not list regulations as a challenge given that it does not list the tax burden as a challenge.

4.52 According to U.S. Census Bureau figures, 35.3% of all Americans are in the 0–24 age bracket, 14.2% are in the 25–34 age bracket, 16.0% are in the 35–44 age bracket, and 34.5 are in the 45 and older age bracket. A study by Jupiter Media Metrix determined that Americans use their leisure time in different ways according to age. For example, of those who are in the 45 and older age bracket, 39% read a book or a magazine more than 10 hours per week. Of those who are in the 0–24 age bracket, only 11% read a book or a magazine more than 10 hours per week. The percentage figures for reading a book or a magazine for more than 10 hours per week are 24% for the 25–34 age bracket and 27% the 35–44 age bracket. Suppose an American is randomly selected and it is determined that he or she reads a book or a magazine more than 10 hours per week. Revise the probabilities that he or she is in any given age category. Using these figures, what is the overall percentage of the U.S. population that reads a book or a magazine more than 10 hours per week?

4.53 A retail study by Deloitte revealed that 54% of adults surveyed believed that plastic, noncompostable shopping bags should be banned. Suppose 41% of adults regularly recycle aluminum cans and believe that plastic, noncompostable shopping bags should be banned. In addition, suppose that 60% of adults who do not believe that plastic, noncompostable shopping bags should be banned do recycle. If an adult is randomly selected,

a. What is the probability that the adult recycles and does not believe that plastic, noncompostable shopping bags should be banned?

b. What is the probability that the adult does recycle?

c. What is the probability that the adult does recycle or does believe that plastic, noncompostable shopping bags should be banned?

d. What is the probability that the adult does not recycle or does not believe that plastic, noncompostable shopping bags should be banned?

e. What is the probability that the adult does not believe that plastic, noncompostable shopping bags should be banned given that the adult does recycle?

ANALYZING THE DATABASES

see www.wiley.com/college/black

1. In the Manufacturing database, what is the probability that a randomly selected SIC Code industry is in industry group 13? What is the probability that a randomly selected SIC Code industry has a value of industry shipments of 4 (see Chapter 1 for coding)? What is the probability that a randomly selected SIC Code industry is in industry group 13 and has a value of industry shipments of 2? What is the probability that a randomly selected SIC Code industry is in industry group 13 or has a value of industry shipments of 2? What is the probability that a randomly selected SIC Code industry neither is in industry group 13 nor has a value of industry shipments of 2?

2. Use the Hospital database. Construct a raw values matrix for region and for type of control. You should have a 7×4 matrix. Using this matrix, answer the following questions. (Refer to Chapter 1 for category members.) What is the probability that a randomly selected hospital is in the Midwest if the hospital is known to be for-profit? If the hospital is known to be in the South, what is the probability that it is a government, nonfederal hospital? What is the probability that a hospital is in the Rocky Mountain region or a not-for-profit, nongovernment hospital? What is the probability that a hospital is a for-profit hospital located in California?

CASE

COLGATE-PALMOLIVE MAKES A "TOTAL" EFFORT

In the mid-1990s, Colgate-Palmolive developed a new toothpaste for the U.S. market, Colgate Total, with an antibacterial ingredient that was already being successfully sold overseas. However, the word *antibacterial* was not allowed for such products by the Food and Drug Administration rules. So Colgate-Palmolive had to come up with another way of marketing this and other features of their new toothpaste to U.S. consumers. Market researchers told Colgate-Palmolive that consumers were weary of trying to discern among the different advantages of various toothpaste brands and wanted simplification in their shopping lives. In response, the name "Total" was given to the product in the United States: The one word would convey that the toothpaste is the "total" package of various benefits.

Young & Rubicam developed several commercials illustrating Total's benefits and tested the commercials with focus groups. One commercial touting Total's long-lasting benefits was particularly successful. Meanwhile, in 1997, Colgate-Palmolive received FDA approval for Total, five years after the company had applied for it. The product was launched in the United States in January of 1998 using commercials that were designed from the more successful ideas of the focus group tests. Total was introduced with a $100 million advertising campaign. Ten months later, 21% of all United States households had purchased Total for the first time. During this same time period, 43% of those who initially tried Total purchased it again. A year after its release, Total was the number one toothpaste in the United States. Total is advertised as not just a toothpaste but as a protective shield that protects you for a full range of oral health problems for up to 12 hours. Total is now offered in a variety of forms, including Colgate Total Advanced Whitening, Colgate Total Advanced Clean, Colgate Total Advanced Fresh Gel, Colgate Total Clean Mint Paste, Colgate Total Whitening Paste, Colgate Total Whitening Gel, Colgate Total Plus Whitening Liquid, and Colgate Total Mint Stripe Gel. In the United States, market share for Colgate Total toothpaste was 16.2% in the second quarter of 2008, which was its highest quarterly share ever.

Discussion

1. What probabilities are given in this case? Use these probabilities and the probability laws to determine what percentage of U.S. households purchased Total at least twice in the first 10 months of its release.
2. Is age category independent of willingness to try new products? According to the U.S. Census Bureau, approximately 20% of all Americans are in the 45–64 age category. Suppose 24% of the consumers who purchased Total for the first time during the initial 10-month period were in the 45–64 age category. Use this information to determine whether age is independent of the initial purchase of Total during the introductory time period. Explain your answer.
3. Using the probabilities given in Question 2, calculate the probability that a randomly selected U.S. consumer is either in the 45–64 age category or purchased Total during the initial 10-month period. What is the probability that a randomly selected person purchased Total in the first 10 months given that the person is in the 45–64 age category?
4. Suppose 32% of all toothpaste consumers in the United States saw the Total commercials. Of those who saw the commercials, 40% purchased Total at least once in the first 10 months of its introduction. Of those who did not see the commercials, 12.06% purchased Total at least once in the first 10 months of its introduction. Suppose a toothpaste consumer is randomly selected and it is learned that they purchased Total during the first 10 months of its introduction. Revise the probability that this person saw the Total commercials and the probability that the person did not see the Total commercials.

Source: Colgate-Palmolive's home page at http://www.colgate.com/app/Colgate/US/HomePage.cvsp, Total's homepage at http://www.colgate.com/app/ColgateTotal/US/EN/Products.cvsp, and at answers.com found at: http://www.answers.com/topic/colgate-palmolive-company, 2008