Chapter 1

Introduction to Statistics and Probability

1.1 Overview: Statistical Inference, Samples, Populations, and the Role of Probability

Beginning in the 1980s and continuing into the 21st century, a great deal of attention has been focused on *improvement of quality* in American industry. Much has been said and written about the Japanese "industrial miracle," which began in the middle of the 20th century. The Japanese were able to succeed where we and other countries had failed—namely, to create an atmosphere that allows the production of high-quality products. Much of the success of the Japanese has been attributed to the use of *statistical methods* and statistical thinking among management personnel.

Use of Scientific Data

The use of statistical methods in manufacturing, development of food products, computer software, energy sources, pharmaceuticals, and many other areas involves the gathering of information or **scientific data**. Of course, the gathering of data is nothing new. It has been done for well over a thousand years. Data have been collected, summarized, reported, and stored for perusal. However, there is a profound distinction between collection of scientific information and **inferential statistics**. It is the latter that has received rightful attention in recent decades.

The offspring of inferential statistics has been a large "toolbox" of statistical methods employed by statistical practitioners. These statistical methods are designed to contribute to the process of making scientific judgments in the face of **uncertainty** and **variation**. The product density of a particular material from a manufacturing process will not always be the same. Indeed, if the process involved is a batch process rather than continuous, there will be not only variation in material density among the batches that come off the line (batch-to-batch variation), but also within-batch variation. Statistical methods are used to analyze data from a process such as this one in order to gain more sense of where in the process changes may be made to improve the **quality** of the process. In this process, quality may well be defined in relation to closeness to a target density value in harmony with *what portion of the time* this closeness criterion is met. An engineer may be concerned with a specific instrument that is used to measure sulfur monoxide in the air during pollution studies. If the engineer has doubts about the effectiveness of the instrument, there are two **sources of variation** that must be dealt with. The first is the variation in sulfur monoxide values that are found at the same locale on the same day. The second is the variation between values observed and the **true** amount of sulfur monoxide that is in the air at the time. If either of these two sources of variation is exceedingly large (according to some standard set by the engineer), the instrument may need to be replaced. In a biomedical study of a new drug that reduces hypertension, 85% of patients experienced relief, while it is generally recognized that the current drug, or "old" drug, brings relief to 80% of patients that have chronic hypertension. However, the new drug is more expensive to make and may result in certain side effects. Should the new drug be adopted? This is a problem that is encountered (often with much more complexity) frequently by pharmaceutical firms in conjunction with the FDA (Federal Drug Administration). Again, the consideration of variation needs to be taken into account. The "85%" value is based on a certain number of patients chosen for the study. Perhaps if the study were repeated with new patients the observed number of "successes" would be 75%! It is the natural variation from study to study that must be taken into account in the decision process. Clearly this variation is important, since variation from patient to patient is endemic to the problem.

Variability in Scientific Data

In the problems discussed above the statistical methods used involve dealing with variability, and in each case the variability to be studied is that encountered in scientific data. If the observed product density in the process were always the same and were always on target, there would be no need for statistical methods. If the device for measuring sulfur monoxide always gives the same value and the value is accurate (i.e., it is correct), no statistical analysis is needed. If there were no patient-to-patient variability inherent in the response to the drug (i.e., it either always brings relief or not), life would be simple for scientists in the pharmaceutical firms and FDA and no statistician would be needed in the decision process. Statistics researchers have produced an enormous number of analytical methods that allow for analysis of data from systems like those described above. This reflects the true nature of the science that we call inferential statistics, namely, using techniques that allow us to go beyond merely reporting data to drawing conclusions (or inferences) about the scientific system. Statisticians make use of fundamental laws of probability and statistical inference to draw conclusions about scientific systems. Information is gathered in the form of **samples**, or collections of **observations**. The process of sampling will be introduced in this chapter, and the discussion continues throughout the entire book.

Samples are collected from **populations**, which are collections of all individuals or individual items of a particular type. At times a population signifies a scientific system. For example, a manufacturer of computer boards may wish to eliminate defects. A sampling process may involve collecting information on 50 computer boards sampled randomly from the process. Here, the population is all computer boards manufactured by the firm over a specific period of time. If an improvement is made in the computer board process and a second sample of boards is collected, any conclusions drawn regarding the effectiveness of the change in process should extend to the entire population of computer boards produced under the "improved process." In a drug experiment, a sample of patients is taken and each is given a specific drug to reduce blood pressure. The interest is focused on drawing conclusions about the population of those who suffer from hypertension.

Often, it is very important to collect scientific data in a systematic way, with planning being high on the agenda. At times the planning is, by necessity, quite limited. We often focus only on certain properties or characteristics of the items or objects in the population. Each characteristic has particular engineering or, say, biological importance to the "customer," the scientist or engineer who seeks to learn about the population. For example, in one of the illustrations above the quality of the process had to do with the product density of the output of a process. An engineer may need to study the effect of process conditions, temperature, humidity, amount of a particular ingredient, and so on. He or she can systematically move these **factors** to whatever levels are suggested according to whatever prescription or **experimental design** is desired. However, a forest scientist who is interested in a study of factors that influence wood density in a certain kind of tree cannot necessarily design an experiment. This case may require an **observational study** in which data are collected in the field but **factor levels** can not be preselected. Both of these types of studies lend themselves to methods of statistical inference. In the former, the quality of the inferences will depend on proper planning of the experiment. In the latter, the scientist is at the mercy of what can be gathered. For example, it is sad if an agronomist is interested in studying the effect of rainfall on plant yield and the data are gathered during a drought.

The importance of statistical thinking by managers and the use of statistical inference by scientific personnel is widely acknowledged. Research scientists gain much from scientific data. Data provide understanding of scientific phenomena. Product and process engineers learn a great deal in their off-line efforts to improve the process. They also gain valuable insight by gathering production data (on-line monitoring) on a regular basis. This allows them to determine necessary modifications in order to keep the process at a desired level of quality.

There are times when a scientific practitioner wishes only to gain some sort of summary of a set of data represented in the sample. In other words, inferential statistics is not required. Rather, a set of single-number statistics or **descriptive statistics** is helpful. These numbers give a sense of center of the location of the data, variability in the data, and the general nature of the distribution of observations in the sample. Though no specific statistical methods leading to **statistical inference** are incorporated, much can be learned. At times, descriptive statistics are accompanied by graphics. Modern statistical software packages allow for computation of **means**, **medians**, **standard deviations**, and other single-number statistics as well as production of graphs that show a "footprint" of the nature of the sample, including histograms, stem-and-leaf plots, scatter plots, dot plots, and box plots.

The Role of Probability

From this chapter to Chapter 3, we deal with fundamental notions of probability. A thorough grounding in these concepts allows the reader to have a better understanding of statistical inference. Without some formalism of probability theory, the student cannot appreciate the true interpretation from data analysis through modern statistical methods. It is quite natural to study probability prior to studying statistical inference. Elements of probability allow us to quantify the strength or "confidence" in our conclusions. In this sense, concepts in probability form a major component that supplements statistical methods and helps us gauge the strength of the statistical inference. The discipline of probability, then, provides the transition between descriptive statistics and inferential methods. Elements of probability allow the conclusion to be put into the language that the science or engineering practitioners require. An example follows that will enable the reader to understand the notion of a P-value, which often provides the "bottom line" in the interpretation of results from the use of statistical methods.

Example 1.1: Suppose that an engineer encounters data from a manufacturing process in which 100 items are sampled and 10 are found to be defective. It is expected and anticipated that occasionally there will be defective items. Obviously these 100 items represent the sample. However, it has been determined that in the long run, the company can only tolerate 5% defective in the process. Now, the elements of probability allow the engineer to determine how conclusive the sample information is regarding the nature of the process. In this case, the **population** conceptually represents all possible items from the process. Suppose we learn that *if the process is acceptable*, that is, if it does produce items no more than 5% of which are defective, there is a probability of 0.0282 of obtaining 10 or more defective items in a random sample of 100 items from the process. This small probability suggests that the process does, indeed, have a long-run rate of defective items that exceeds 5%. In other words, under the condition of an acceptable process, the sample information obtained would rarely occur. However, it did occur! Clearly, though, it would occur with a much higher probability if the process defective rate exceeded 5% by a significant amount.

From this example it becomes clear that the elements of probability aid in the translation of sample information into something conclusive or inconclusive about the scientific system. In fact, what was learned likely is alarming information to the engineer or manager. Statistical methods, which we will actually detail in Chapter 6, produced a P-value of 0.0282. The result suggests that the process **very likely is not acceptable**. The concept of a ***P*-value** is dealt with at length in succeeding chapters. The example that follows provides a second illustration.

Example 1.2: Often the nature of the scientific study will dictate the role that probability and deductive reasoning play in statistical inference. Exercise 5.28 on page 221 provides data associated with a study conducted at Virginia Tech on the development of a relationship between the roots of trees and the action of a fungus. Minerals are transferred from the fungus to the trees and sugars from the trees to the fungus. Two samples of 10 northern red oak seedlings were planted in a greenhouse, one containing seedlings treated with nitrogen and the other containing seedlings with no nitrogen. All other environmental conditions were held constant. All seedlings contained the fungus *Pisolithus tinctorus*. More details are supplied in Chapter 5. The stem weights in grams were recorded after the end of 140 days. The data are given in Table 1.1.

Table 1.1: Data Set for Example 1.2

No Nitrogen	**Nitrogen**
0.32	0.26
0.53	0.43
0.28	0.47
0.37	0.49
0.47	0.52
0.43	0.75
0.36	0.79
0.42	0.86
0.38	0.62
0.43	0.46

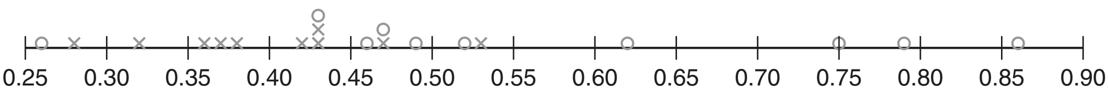

Figure 1.1: A dot plot of stem weight data.

In this example there are two samples from two **separate populations**. The purpose of the experiment is to determine if the use of nitrogen has an influence on the growth of the roots. The study is a comparative study (i.e., we seek to compare the two populations with regard to a certain important characteristic). It is instructive to plot the data as shown in the dot plot of Figure 1.1. The $\circ$ values represent the "nitrogen" data and the $\times$ values represent the "no-nitrogen" data.

Notice that the general appearance of the data might suggest to the reader that, on average, the use of nitrogen increases the stem weight. Four nitrogen observations are considerably larger than any of the no-nitrogen observations. Most of the no-nitrogen observations appear to be below the center of the data. The appearance of the data set would seem to indicate that nitrogen is effective. But how can this be quantified? How can all of the apparent visual evidence be summarized in some sense? As in the preceding example, the fundamentals of probability can be used. The conclusions may be summarized in a probability statement or P-value. We will not show here the statistical inference that produces the summary probability. As in Example 1.1, these methods will be discussed in Chapter 6. The issue revolves around the "probability that data like these could be observed" *given that nitrogen has no effect*, in other words, given that both samples were generated from the same population. Suppose that this probability is small, say 0.03. That would certainly be strong evidence that the use of nitrogen does indeed influence (apparently increases) average stem weight of the red oak seedlings. ■

How Do Probability and Statistical Inference Work Together?

It is important for the reader to understand the clear distinction between the discipline of probability, a science in its own right, and the discipline of inferential statistics. As we have already indicated, the use or application of concepts in probability allows real-life interpretation of the results of statistical inference. As a result, it can be said that statistical inference makes use of concepts in probability. One can glean from the two examples above that the sample information is made available to the analyst and, with the aid of statistical methods and elements of probability, conclusions are drawn about some feature of the population (the process does not appear to be acceptable in Example 1.1, and nitrogen does appear to influence average stem weights in Example 1.2). Thus for a statistical problem, **the sample along with inferential statistics allows us to draw conclusions about the population, with inferential statistics making clear use of elements of probability**. This reasoning is *inductive* in nature. Now as we move into Section 1.4 and beyond, the reader will note that, unlike what we do in our two examples here, we will not focus on solving statistical problems. Many examples will be given in which no sample is involved. There will be a population clearly described with all features of the population known. Then questions of importance will focus on the nature of data that might hypothetically be drawn from the population. Thus, one can say that **elements in probability allow us to draw conclusions about characteristics of hypothetical data taken from the population, based on known features of the population**. This type of reasoning is *deductive* in nature. Figure 1.2 shows the fundamental relationship between probability and inferential statistics.

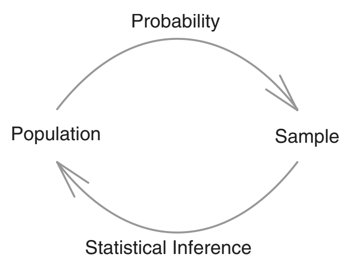

Figure 1.2: Fundamental relationship between probability and inferential statistics.

Now, in the grand scheme of things, which is more important, the field of probability or the field of statistics? They are both very important and clearly are complementary. The only certainty concerning the pedagogy of the two disciplines lies in the fact that if statistics is to be taught at more than merely a "cookbook" level, then the discipline of probability must be taught first. This rule stems from the fact that nothing can be learned about a population from a sample until the analyst learns the rudiments of uncertainty in that sample. For example, consider Example 1.1. The question centers around whether or not the population, defined by the process, is no more than 5% defective. In other words, the conjecture is that **on the average** 5 out of 100 items are defective. Now, the sample contains 100 items and 10 are defective. Does this support the conjecture or refute it? On the surface it would appear to be a refutation of the conjecture because 10 out of 100 seem to be "a bit much." But without elements of probability, how do we know? Only through the study of material in future chapters will we learn the conditions under which the process is acceptable (5% defective). The probability of obtaining 10 or more defective items in a sample of 100 is 0.0282.

We have given two examples where the elements of probability provide a summary that the scientist or engineer can use as evidence on which to build a decision. The bridge between the data and the conclusion is, of course, based on foundations of statistical inference, distribution theory, and sampling distributions discussed in future chapters.

1.2 Sampling Procedures; Collection of Data

In Section 1.1 we discussed very briefly the notion of sampling and the sampling process. While sampling appears to be a simple concept, the complexity of the questions that must be answered about the population or populations necessitates that the sampling process be very complex at times. While the notion of sampling is discussed in a technical way in Chapter 4, we shall endeavor here to give some common-sense notions of sampling. This is a natural transition to a discussion of the concept of variability.

Simple Random Sampling

The importance of proper sampling revolves around the degree of confidence with which the analyst is able to answer the questions being asked. Let us assume that only a single population exists in the problem. Recall that in Example 1.2 two populations were involved. **Simple random sampling** implies that any particular sample of a specified *sample size* has the same chance of being selected as any other sample of the same size. The term **sample size** simply means the number of elements in the sample. Obviously, a table of random numbers can be utilized in sample selection in many instances. The virtue of simple random sampling is that it aids in the elimination of the problem of having the sample reflect a different (possibly more confined) population than the one about which inferences need to be made. For example, a sample is to be chosen to answer certain questions regarding political preferences in a certain state in the United States. The sample involves the choice of, say, 1000 families, and a survey is to be conducted. Now, suppose it turns out that random sampling is not used. Rather, all or nearly all of the 1000 families chosen live in an urban setting. It is believed that political preferences in rural areas differ from those in urban areas. In other words, the sample drawn actually confined the population and thus the inferences need to be confined to the "limited population," and in this case confining may be undesirable. If, indeed, the inferences need to be made about the state as a whole, the sample of size 1000 described here is often referred to as a **biased sample**.

As we hinted earlier, simple random sampling is not always appropriate. Which alternative approach is used depends on the complexity of the problem. Often, for example, the sampling units are not homogeneous and naturally divide themselves into nonoverlapping groups that are homogeneous. These groups are called *strata*, and a procedure called *stratified random sampling* involves random selection of a sample *within* each stratum. The purpose is to be sure that each of the strata is neither over- nor underrepresented. For example, suppose a sample survey is conducted in order to gather preliminary opinions regarding a bond referendum that is being considered in a certain city. The city is subdivided into several ethnic groups which represent natural strata. In order not to disregard or overrepresent any group, separate random samples of families could be chosen from each group.

Experimental Design

The concept of randomness or random assignment plays a huge role in the area of **experimental design**, which was introduced very briefly in Section 1.1 and is an important staple in almost any area of engineering or experimental science. This will also be discussed at length in Chapter 8. However, it is instructive to give a brief presentation here in the context of random sampling. A set of so-called **treatments** or **treatment combinations** becomes the populations to be studied or compared in some sense. An example is the nitrogen versus no-nitrogen treatments in Example 1.2. Another simple example would be placebo versus active drug, or in a corrosion fatigue study we might have treatment combinations that involve specimens that are coated or uncoated as well as conditions of low or high humidity to which the specimens are exposed. In fact, there are four treatment or factor combinations (i.e., 4 populations), and many scientific questions may be asked and answered through statistical and inferential methods. Consider first the situation in Example 1.2. There are 20 diseased seedlings involved in the experiment. It is easy to see from the data themselves that the seedlings are different from each other. Within the nitrogen group (or the no-nitrogen group) there is considerable **variability** in the stem weights. This variability is due to what is generally called the **experimental unit**. This is a very important concept in inferential statistics, in fact one whose description will not end in this chapter. The nature of the variability is very important. If it is too large, stemming from a condition of excessive nonhomogeneity in experimental units, the variability will "wash out" any detectable difference between the two populations. Recall that in this case that did not occur.

The dot plot in Figure 1.1 and P-value indicated a clear distinction between these two conditions. What role do those experimental units play in the data-taking process itself? The common-sense and, indeed, quite standard approach is to assign the 20 seedlings or experimental units **randomly to the two treatments or conditions**. In the drug study, we may decide to use a total of 200 available patients, patients that clearly will be different in some sense. They are the experimental units. However, they all may have the same chronic condition for which the drug is a potential treatment. Then in a so-called **completely randomized design**, 100 patients are assigned randomly to the placebo and 100 to the active drug. Again, it is these experimental units within a group or treatment that produce the variability in data results (i.e., variability in the measured result), say blood pressure, or whatever drug efficacy value is important. In the corrosion fatigue study, the experimental units are the specimens that are the subjects of the corrosion.

Why Assign Experimental Units Randomly?

What is the possible negative impact of not randomly assigning experimental units to the treatments or treatment combinations? This is seen most clearly in the case of the drug study. Among the characteristics of the patients that produce variability in the results are age, gender, and weight. Suppose merely by chance the placebo group contains a sample of people that are predominately heavier than those in the treatment group. Perhaps heavier individuals have a tendency to have a higher blood pressure. This clearly biases the result, and indeed, any result obtained through the application of statistical inference may have little to do with the drug and more to do with differences in weights among the two samples of patients.

We should emphasize the attachment of importance to the term **variability**. Excessive variability among experimental units "camouflages" scientific findings. In future sections, we attempt to characterize and quantify measures of variability. In sections that follow, we introduce and discuss specific quantities that can be computed in samples; the quantities give a sense of the nature of the sample with respect to center of location of the data and variability in the data. A discussion of several of these single-number measures serves to provide a preview of what statistical information will be important components of the statistical methods that are used in future chapters. These measures that help characterize the nature of the data set fall into the category of **descriptive statistics**. This material is a prelude to a brief presentation of pictorial and graphical methods that go even further in characterization of the data set. The reader should understand that the statistical methods illustrated here will be used throughout the text. In order to offer the reader a clearer picture of what is involved in experimental design studies, we offer Example 1.3.

Example 1.3: A corrosion study was made in order to determine whether coating an aluminum metal with a corrosion retardation substance reduced the amount of corrosion. The coating is a protectant that is advertised to minimize fatigue damage in this type of material. Also of interest is the influence of humidity on the amount of corrosion. A corrosion measurement can be expressed in thousands of cycles to failure. Two levels of coating, no coating and chemical corrosion coating, were used. In addition, the two relative humidity levels are 20% relative humidity and 80% relative humidity.

The experiment involves four treatment combinations that are listed in the table that follows. There are eight experimental units used, and they are aluminum specimens prepared; two are assigned randomly to each of the four treatment combinations. The data are presented in Table 1.2.

The corrosion data are averages of two specimens. A plot of the averages is pictured in Figure 1.3. A relatively large value of cycles to failure represents a small amount of corrosion. As one might expect, an increase in humidity appears to make the corrosion worse. The use of the chemical corrosion coating procedure appears to reduce corrosion. ■

In this experimental design illustration, the engineer has systematically selected the four treatment combinations. In order to connect this situation to concepts with which the reader has been exposed to this point, it should be assumed that the

Table 1.2: Data for Example 1.3

Coating	Humidity	Average Corrosion in Thousands of Cycles to Failure
Uncoated	20%	975
	80%	350
Chemical Corrosion	20%	1750
	80%	1550

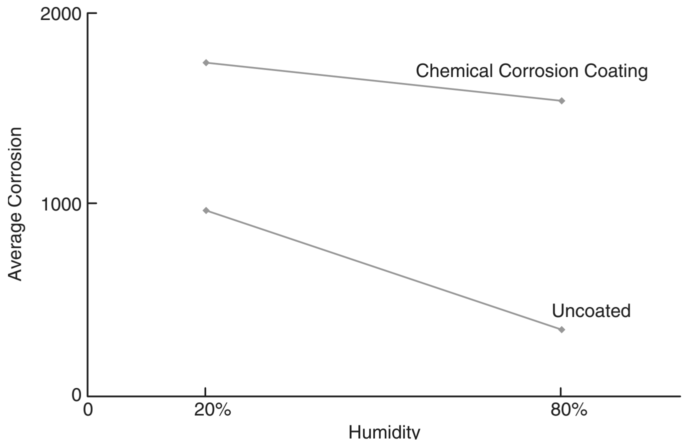

Figure 1.3: Corrosion results for Example 1.3.

conditions representing the four treatment combinations are four separate populations and that the two corrosion values observed for each population are important pieces of information. The importance of the average in capturing and summarizing certain features in the population will be highlighted in Section 4.2. While we might draw conclusions about the role of humidity and the impact of coating the specimens from the figure, we cannot truly evaluate the results from an analytical point of view without taking into account the *variability around* the average. Again, as we indicated earlier, if the two corrosion values for each treatment combination are close together, the picture in Figure 1.3 may be an accurate depiction. But if each corrosion value in the figure is an average of two values that are widely dispersed, then this variability may, indeed, truly "wash away" any information that appears to come through when one observes averages only. The foregoing example illustrates these concepts:

(1) random assignment of treatment combinations (coating, humidity) to experimental units (specimens)

(2) the use of sample averages (average corrosion values) in summarizing sample information

(3) the need for consideration of measures of variability in the analysis of any sample or sets of samples

1.3 Discrete and Continuous Data

Statistical inference through the analysis of observational studies or designed experiments is used in many scientific areas. The data gathered may be **discrete** or **continuous**, depending on the area of application. For example, a chemical engineer may be interested in conducting an experiment that will lead to conditions where yield is maximized. Here, of course, the yield may be in percent or grams/pound, measured on a continuum. On the other hand, a toxicologist conducting a combination drug experiment may encounter data that are binary in nature (i.e., the patient either responds or does not).

Great distinctions are made between discrete and continuous data in the probability theory that allow us to draw statistical inferences. Often applications of statistical inference are found when the data are *count data*. For example, an engineer may be interested in studying the number of radioactive particles passing through a counter in, say, 1 millisecond. Personnel responsible for the efficiency of a port facility may be interested in the properties of the number of oil tankers arriving each day at a certain port city. In Chapter 3, several distinct scenarios, leading to different ways of handling data, are discussed for situations with count data.

Special attention even at this early stage of the textbook should be paid to some details associated with binary data. Applications requiring statistical analysis of binary data are voluminous. Often the measure that is used in the analysis is the *sample proportion*. Obviously the binary situation involves two categories. If there are n units involved in the data and x is defined as the number that fall into category 1, then $n - x$ fall into category 2. Thus, x/n is the sample proportion in category 1, and $1 - x/n$ is the sample proportion in category 2. In the biomedical application, 50 patients may represent the sample units, and if 20 out of 50 experienced an improvement in a stomach ailment (common to all 50) after all were given the drug, then $\frac{20}{50} = 0.4$ is the sample proportion for which the drug was a success and $1 - 0.4 = 0.6$ is the sample proportion for which the drug was not successful. Actually the basic numerical measurement for binary data is generally denoted by either 0 or 1. For example, in our medical example, a successful result is denoted by a 1 and a nonsuccess by a 0. As a result, the sample proportion is actually a sample mean of the ones and zeros. For the successful category,

$$\frac{x_1 + x_2 + \cdots + x_{50}}{50} = \frac{1 + 1 + 0 + \cdots + 0 + 1}{50} = \frac{20}{50} = 0.4.$$

1.4 Probability: Sample Space and Events

Sample Space

In the study of statistics, we are concerned basically with the presentation and interpretation of **chance outcomes** that occur in a planned study or scientific investigation. For example, we may record the number of accidents that occur monthly at the intersection of Driftwood Lane and Royal Oak Drive, hoping to justify the installation of a traffic light; we might classify items coming off an assembly line as "defective" or "nondefective"; or we may be interested in the volume of gas released in a chemical reaction when the concentration of an acid is varied. Hence, the statistician is often dealing with either numerical data, representing counts or measurements, or **categorical data**, which can be classified according to some criterion.

We shall refer to any recording of information, whether it be numerical or categorical, as an **observation**. Thus, the numbers 2, 0, 1, and 2, representing the number of accidents that occurred for each month from January through April during the past year at the intersection of Driftwood Lane and Royal Oak Drive, constitute a set of observations. Similarly, the categorical data N, D, N, N, and D, representing the items found to be defective or nondefective when five items are inspected, are recorded as observations.

Statisticians use the word **experiment** to describe any process that generates a set of data. A simple example of a statistical experiment is the tossing of a coin. In this experiment, there are only two possible outcomes, heads or tails. Another experiment might be the launching of a missile and observing of its velocity at specified times. The opinions of voters concerning a new sales tax can also be considered as observations of an experiment. We are particularly interested in the observations obtained by repeating the experiment several times. In most cases, the outcomes will depend on chance and, therefore, cannot be predicted with certainty. If a chemist runs an analysis several times under the same conditions, he or she will obtain different measurements, indicating an element of chance in the experimental procedure. Even when a coin is tossed repeatedly, we cannot be certain that a given toss will result in a head. However, we know the entire set of possibilities for each toss.

Definition 1.1: The set of all possible outcomes of a statistical experiment is called the **sample space** and is represented by the symbol S.

Each outcome in a sample space is called an **element** or a **member** of the sample space, or simply a **sample point**. If the sample space has a finite number of elements, we may *list* the members separated by commas and enclosed in braces. Thus, the sample space S, of possible outcomes when a coin is flipped, may be written

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

where H and T correspond to heads and tails, respectively.

Example 1.4: Consider the experiment of tossing a die. If we are interested in the number that shows on the top face, the sample space is

$$S_1 = \{1, 2, 3, 4, 5, 6\}.$$

If we are interested only in whether the number is even or odd, the sample space is simply

$$S_2 = \{\text{even}, \text{odd}\}.$$

Example 1.4 illustrates the fact that more than one sample space can be used to describe the outcomes of an experiment. In this case, S_1 provides more information than S_2. If we know which element in S_1 occurs, we can tell which outcome in S_2 occurs; however, a knowledge of what happens in S_2 is of little help in determining which element in S_1 occurs. In general, it is desirable to use the sample space that gives the most information concerning the outcomes of the experiment. In some experiments, it is helpful to list the elements of the sample space systematically by means of a **tree diagram**.

Example 1.5: Suppose that three items are selected at random from a manufacturing process. Each item is inspected and classified defective, D, or nondefective, N. To list the elements of the sample space providing the most information, we construct the tree diagram of Figure 1.4. Now, the various paths along the branches of the tree give the distinct sample points. Starting with the first path, we get the sample point DDD, indicating the possibility that all three items inspected are defective. As we proceed along the other paths, we see that the sample space is

$$S = \{DDD, DDN, DND, DNN, NDD, NDN, NND, NNN\}.$$

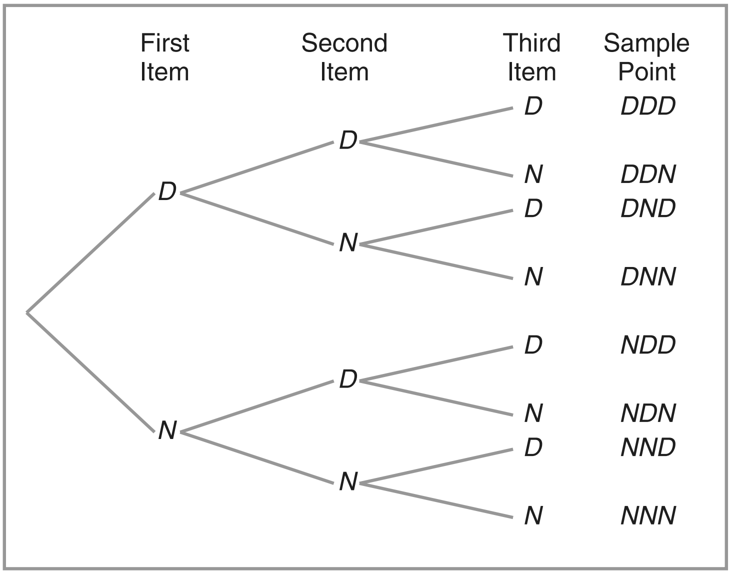

Figure 1.4: Tree diagram for Example 1.5.

Sample spaces with a large or infinite number of sample points are best described by a **statement** or **rule method**. For example, if the possible outcomes of an experiment are the set of cities in the world with a population over 1 million, our sample space is written

$$S = \{x \mid x \text{ is a city with a population over 1 million}\},$$

which reads "S is the set of all x such that x is a city with a population over 1 million." The vertical bar is read "such that." Similarly, if S is the set of all points

(x, y) on the boundary or the interior of a circle of radius 2 with center at the origin, we write the **rule**

$$S = \{(x, y) \mid x^2 + y^2 \leq 4\}.$$

Whether we describe the sample space by the rule method or by listing the elements will depend on the specific problem at hand. The rule method has practical advantages, particularly for many experiments where listing becomes a tedious chore.

Consider the situation of Example 1.5 in which items from a manufacturing process are either D, defective, or N, nondefective. There are many important statistical procedures called sampling plans that determine whether or not a "lot" of items is considered satisfactory. One such plan involves sampling until k defectives are observed. Suppose the experiment is to sample items randomly until one defective item is observed. The sample space for this case is

$$S = \{D, ND, NND, NNND, \dots\}.$$

Events

For any given experiment, we may be interested in the occurrence of certain **events** rather than in the occurrence of a specific element in the sample space. For instance, we may be interested in the event A that the outcome when a die is tossed is divisible by 3. This will occur if the outcome is an element of the subset $A = \{3, 6\}$ of the sample space S_1 in Example 1.4. As a further illustration, we may be interested in the event B that the number of defectives is greater than 1 in Example 1.5. This will occur if the outcome is an element of the subset

$$B = \{DDN, DND, NDD, DDD\}$$

of the sample space S.

To each event we assign a collection of sample points, which constitute a subset of the sample space. That subset represents all of the elements for which the event is true.

Definition 1.2: An **event** is a subset of a sample space.

Example 1.6: Given the sample space $S = \{t \mid t \geq 0\}$, where t is the life in years of a certain electronic component, then the event A that the component fails before the end of the fifth year is the subset $A = \{t \mid 0 \leq t < 5\}$.

It is conceivable that an event may be a subset that includes the entire sample space S or a subset of S called the **null set** and denoted by the symbol ϕ, which contains no elements at all. For instance, if we let A be the event of detecting a microscopic organism with the naked eye in a biological experiment, then $A = \phi$. Also, if

$$B = \{x \mid x \text{ is an even factor of } 7\},$$

then B must be the null set, since the only possible factors of 7 are the odd numbers 1 and 7.

Consider an experiment where the smoking habits of the employees of a manufacturing firm are recorded. A possible sample space might classify an individual as a nonsmoker, a light smoker, a moderate smoker, or a heavy smoker. Let the subset of smokers be some event. Then all the nonsmokers correspond to a different event, also a subset of S, which is called the **complement** of the set of smokers.

Definition 1.3: The **complement** of an event A with respect to S is the subset of all elements of S that are not in A. We denote the complement of A by the symbol A'.

Example 1.7: Let R be the event that a red card is selected from an ordinary deck of 52 playing cards, and let S be the entire deck. Then R' is the event that the card selected from the deck is not a red card but a black card.

Example 1.8: Consider the sample space

$$S = \{\text{book, cell phone, mp3, paper, stationery, laptop}\}.$$

Let $A = \{\text{book, stationery, laptop, paper}\}$. Then the complement of A is $A' = \{\text{cell phone, mp3}\}$.

We now consider certain operations with events that will result in the formation of new events. These new events will be subsets of the same sample space as the given events. Suppose that A and B are two events associated with an experiment. In other words, A and B are subsets of the same sample space S. For example, in the tossing of a die we might let A be the event that an even number occurs and B the event that a number greater than 3 shows. Then the subsets $A = \{2, 4, 6\}$ and $B = \{4, 5, 6\}$ are subsets of the same sample space

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

Note that *both* A and B will occur on a given toss if the outcome is an element of the subset $\{4, 6\}$, which is just the **intersection** of A and B.

Definition 1.4: The **intersection** of two events A and B, denoted by the symbol $A \cap B$, is the event containing all elements that are common to A and B.

Example 1.9: Let E be the event that a person selected at random in a classroom is majoring in engineering, and let F be the event that the person is female. Then $E \cap F$ is the event of all female engineering students in the classroom.

Example 1.10: Let $V = \{a, e, i, o, u\}$ and $C = \{l, r, s, t\}$; then it follows that $V \cap C = \phi$. That is, V and C have no elements in common and, therefore, cannot both simultaneously occur.

For certain statistical experiments it is by no means unusual to define two events, A and B, that cannot both occur simultaneously. The events A and B are then said to be **mutually exclusive**. Stated more formally, we have the following definition:

Definition 1.5: Two events A and B are **mutually exclusive**, or **disjoint**, if $A \cap B = \phi$, that is, if A and B have no elements in common.

Example 1.11: A cable television company offers programs on eight different channels, three of which are affiliated with ABC, two with NBC, and one with CBS. The other two are an educational channel and the ESPN sports channel. Suppose that a person subscribing to this service turns on a television set without first selecting the channel. Let A be the event that the program belongs to the NBC network and B the event that it belongs to the CBS network. Since a television program cannot belong to more than one network, the events A and B have no programs in common. Therefore, the intersection $A \cap B$ contains no programs, and consequently the events A and B are mutually exclusive.

Often one is interested in the occurrence of at least one of two events associated with an experiment. Thus, in the die-tossing experiment, if

$$A = \{2, 4, 6\} \text{ and } B = \{4, 5, 6\},$$

we might be interested in either A or B occurring or both A and B occurring. Such an event, called the **union** of A and B, will occur if the outcome is an element of the subset $\{2, 4, 5, 6\}$.

Definition 1.6: The **union** of the two events A and B, denoted by the symbol $A \cup B$, is the event containing all the elements that belong to A or B or both.

Example 1.12: Let $A = \{a, b, c\}$ and $B = \{b, c, d, e\}$; then $A \cup B = \{a, b, c, d, e\}$.

Example 1.13: Let P be the event that an employee selected at random from an oil drilling company smokes cigarettes. Let Q be the event that the employee selected drinks alcoholic beverages. Then the event $P \cup Q$ is the set of all employees who either drink or smoke or do both.

Example 1.14: If $M = \{x \mid 3 < x < 9\}$ and $N = \{y \mid 5 < y < 12\}$, then

$$M \cup N = \{z \mid 3 < z < 12\}.$$

The relationship between events and the corresponding sample space can be illustrated graphically by means of **Venn diagrams**. In a Venn diagram we let the sample space be a rectangle and represent events by circles drawn inside the rectangle. Thus, in Figure 1.5, we see that

$$\begin{aligned} A \cap B &= \text{ regions 1 and 2,} \\ B \cap C &= \text{ regions 1 and 3,} \\ A \cup C &= \text{ regions 1, 2, 3, 4, 5, and 7,} \\ B' \cap A &= \text{ regions 4 and 7,} \\ A \cap B \cap C &= \text{ region 1,} \\ (A \cup B) \cap C' &= \text{ regions 2, 6, and 7,} \end{aligned}$$

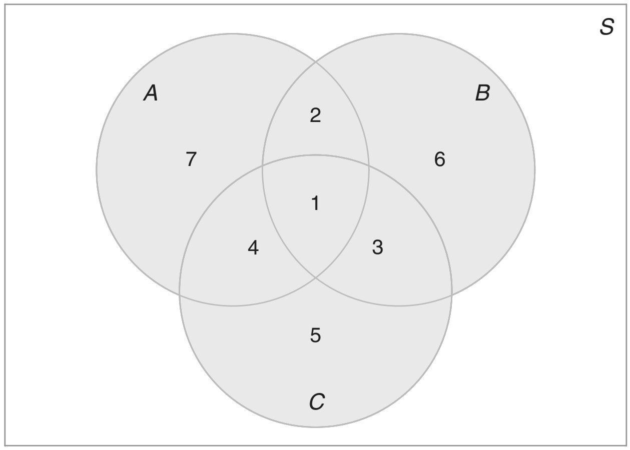

Figure 1.5: Events represented by various regions.

and so forth.

In Figure 1.6, we see that events A, B, and C are all subsets of the sample space S. It is also clear that event B is a subset of event A; event $B \cap C$ has no elements and hence B and C are mutually exclusive; event $A \cap C$ has at least one element; and event $A \cup B = A$. Figure 1.6 might, therefore, depict a situation where we select a card at random from an ordinary deck of 52 playing cards and observe whether the following events occur:

A: the card is red,

B: the card is the jack, queen, or king of diamonds,

C: the card is an ace.

Clearly, the event $A \cap C$ consists of only the two red aces.

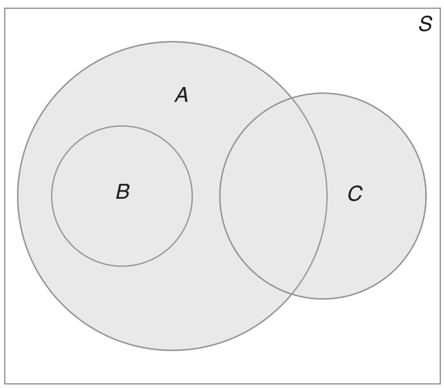

Figure 1.6: Events of the sample space S.

Several results that follow from the foregoing definitions, which may easily be verified by means of Venn diagrams, are as follows:

1. $A \cap \phi = \phi$.
2. $A \cup \phi = A$.
3. $A \cap A' = \phi$.
4. $A \cup A' = S$.
5. $S' = \phi$.
6. $\phi' = S$.
7. $(A')' = A$.
8. $(A \cap B)' = A' \cup B'$.
9. $(A \cup B)' = A' \cap B'$.

Exercises

1.1 List the elements of each of the following sample spaces:

(a) the set of integers between 1 and 50 divisible by 8;

(b) the set $S = \{x \mid x^2 + 4x - 5 = 0\}$;

(c) the set of outcomes when a coin is tossed until a tail or three heads appear;

(d) the set $S = \{x \mid x \text{ is a continent}\}$;

(e) the set $S = \{x \mid 2x - 4 \geq 0 \text{ and } x < 1\}$.

1.2 Use the rule method to describe the sample space S consisting of all points in the first quadrant inside a circle of radius 3 with center at the origin.

1.3 Which of the following events are equal?

(a) $A = \{1, 3\}$;

(b) $B = \{x \mid x \text{ is a number on a die}\}$;

(c) $C = \{x \mid x^2 - 4x + 3 = 0\}$;

(d) $D = \{x \mid x$ is the number of heads when six coins are tossed$\}$.

1.4 Two jurors are selected from 4 alternates to serve at a murder trial. Using the notation A_1A_3, for example, to denote the simple event that alternates 1 and 3 are selected, list the 6 elements of the sample space S.

1.5 An experiment consists of tossing a die and then flipping a coin once if the number on the die is even. If the number on the die is odd, the coin is flipped twice. Using the notation $4H$, for example, to denote the outcome that the die comes up 4 and then the coin comes up heads, and $3HT$ to denote the outcome that the die comes up 3 followed by a head and then a tail on the coin, construct a tree diagram to show the 18 elements of the sample space S.

1.6 For the sample space of Exercise 1.5,

(a) list the elements corresponding to the event A that a number less than 3 occurs on the die;

(b) list the elements corresponding to the event B that two tails occur;

(c) list the elements corresponding to the event A';

(d) list the elements corresponding to the event $A' \cap B$;

(e) list the elements corresponding to the event $A \cup B$.

1.7 The resumés of two male applicants for a college teaching position in chemistry are placed in the same file as the resumés of two female applicants. Two positions become available, and the first, at the rank of assistant professor, is filled by selecting one of the four applicants at random. The second position, at the rank of instructor, is then filled by selecting at random one of the remaining three applicants. Using the notation M_2F_1, for example, to denote the simple event that the first position is filled by the second male applicant and the second position is then filled by the first female applicant,

(a) list the elements of a sample space S;

(b) list the elements of S corresponding to event A that the position of assistant professor is filled by a male applicant;

(c) list the elements of S corresponding to event B that exactly one of the two positions is filled by a male applicant;

(d) list the elements of S corresponding to event C that neither position is filled by a male applicant;

(e) list the elements of S corresponding to the event $A \cap B$;

(f) list the elements of S corresponding to the event $A \cup C$;

(g) construct a Venn diagram to illustrate the intersections and unions of the events A, B, and C.

1.8 An engineering firm is hired to determine if certain waterways in Virginia are safe for fishing. Samples are taken from three rivers.

(a) List the elements of a sample space S, using the letters F for safe to fish and N for not safe to fish.

(b) List the elements of S corresponding to event E

that at least two of the rivers are safe for fishing.

(c) Define an event that has as its elements the points

$$\{FFF, NFF, FFN, NFN\}.$$

1.9 Construct a Venn diagram to illustrate the possible intersections and unions for the following events relative to the sample space consisting of all automobiles made in the United States.

F: Four door, S: Sun roof, P: Power steering.

1.10 Exercise and diet are being studied as possible substitutes for medication to lower blood pressure. Three groups of subjects will be used to study the effect of exercise. Group 1 is sedentary, while group 2 walks and group 3 swims for 1 hour a day. Half of each of the three exercise groups will be on a salt-free diet. An additional group of subjects will not exercise or restrict their salt, but will take the standard medication. Use Z for sedentary, W for walker, S for swimmer, Y for salt, N for no salt, M for medication, and F for medication free.

(a) Show all of the elements of the sample space S.

(b) Given that A is the set of nonmedicated subjects and B is the set of walkers, list the elements of $A \cup B$.

(c) List the elements of $A \cap B$.

1.11 If $S = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ and $A = \{0, 2, 4, 6, 8\}$, $B = \{1, 3, 5, 7, 9\}$, $C = \{2, 3, 4, 5\}$, and $D = \{1, 6, 7\}$, list the elements of the sets corresponding to the following events:

(a) $A \cup C$;

(b) $A \cap B$;

(c) C';

(d) $(C' \cap D) \cup B$;

(e) $(S \cap C)'$;

(f) $A \cap C \cap D'$.

1.12 If $S = \{x \mid 0 < x < 12\}$, $M = \{x \mid 1 < x < 9\}$, and $N = \{x \mid 0 < x < 5\}$, find

(a) $M \cup N$;

(b) $M \cap N$;

(c) $M' \cap N'$.

1.13 Let A, B, and C be events relative to the sample space S. Using Venn diagrams, shade the areas representing the following events:

(a) $(A \cap B)'$;

(b) $(A \cup B)'$;

(c) $(A \cap C) \cup B$.

1.14 Which of the following pairs of events are mutually exclusive?

(a) A golfer scoring the lowest 18-hole round in a 72-hole tournament and losing the tournament.

(b) A poker player getting a flush (all cards in the same suit) and 3 of a kind on the same 5-card hand.

(c) A mother giving birth to a baby girl and a set of twin daughters on the same day.

(d) A chess player losing the last game and winning the match.

1.15 Suppose that a family is leaving on a summer vacation in their camper and that M is the event that they will experience mechanical problems, T is the event that they will receive a ticket for committing a traffic violation, and V is the event that they will arrive at a campsite with no vacancies. Referring to the Venn diagram of Figure 1.7, state in words the events represented by the following regions:

(a) region 5;

(b) region 3;

(c) regions 1 and 2 together;

(d) regions 4 and 7 together;

(e) regions 3, 6, 7, and 8 together.

1.16 Referring to Exercise 1.15 and the Venn diagram of Figure 1.7, list the numbers of the regions that represent the following events:

(a) The family will experience no mechanical problems and will not receive a ticket for a traffic violation but will arrive at a campsite with no vacancies.

(b) The family will experience both mechanical problems and trouble in locating a campsite with a vacancy but will not receive a ticket for a traffic violation.

(c) The family will either have mechanical trouble or arrive at a campsite with no vacancies but will not receive a ticket for a traffic violation.

(d) The family will not arrive at a campsite with no vacancies.

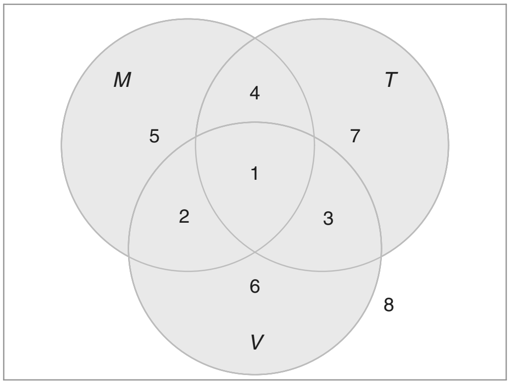

Figure 1.7: Venn diagram for Exercises 1.15 and 1.16.

1.5 Counting Sample Points

Frequently, we are interested in a sample space that contains as elements all possible orders or arrangements of a group of objects. For example, we may want to know how many different arrangements are possible for sitting 6 people around a table, or we may ask how many different orders are possible for drawing 2 lottery tickets from a total of 20. The different arrangements are called **permutations**.

Definition 1.7: A **permutation** is an arrangement of all or part of a set of objects.

Consider the three letters a, b, and c. The possible permutations are abc, acb, bac, bca, cab, and cba. Thus, we see that there are 6 distinct arrangements.

Theorem 1.1: The number of permutations of n objects is $n!$.

The number of permutations of the four letters a, b, c, and d will be $4! = 24$. Now consider the number of permutations that are possible by taking two letters at a time from four. These would be ab, ac, ad, ba, bc, bd, ca, cb, cd, da, db, and dc. Consider that we have two positions to fill, with $n_1 = 4$ choices for the first and then $n_2 = 3$ choices for the second, for a total of

$$n_1 n_2 = (4)(3) = 12$$

permutations. In general, n distinct objects taken r at a time can be arranged in

$$n(n-1)(n-2)\cdots(n-r+1)$$

ways. We represent this product by the symbol

$${}_nP_r = \frac{n!}{(n-r)!}.$$

As a result, we have the theorem that follows.

Theorem 1.2: The number of permutations of n distinct objects taken r at a time is

$$_nP_r = \frac{n!}{(n-r)!}.$$

Example 1.15: In one year, three awards (research, teaching, and service) will be given to a class of 25 graduate students in a statistics department. If each student can receive at most one award, how many possible selections are there?

***Solution*:** Since the awards are distinguishable, it is a permutation problem. The total number of sample points is

$$_{25}P_3 = \frac{25!}{(25-3)!} = \frac{25!}{22!} = (25)(24)(23) = 13{,}800.$$

Example 1.16: A president and a treasurer are to be chosen from a student club consisting of 50 people. How many different choices of officers are possible if

(a) there are no restrictions;

(b) A will serve only if he is president;

(c) B and C will serve together or not at all;

(d) D and E will not serve together?

***Solution*:** (a) The total number of choices of officers, without any restrictions, is

$$_{50}P_2 = \frac{50!}{48!} = (50)(49) = 2450.$$

(b) Since A will serve only if he is president, we have two situations here: (i) A is selected as the president, which yields 49 possible outcomes for the treasurer's position, or (ii) officers are selected from the remaining 49 people without A, which has the number of choices $_{49}P_2 = (49)(48) = 2352$. Therefore, the total number of choices is $49 + 2352 = 2401$.

(c) The number of selections when B and C serve together is 2. The number of selections when both B and C are not chosen is $_{48}P_2 = 2256$. Therefore, the total number of choices in this situation is $2 + 2256 = 2258$.

(d) The number of selections when D serves as an officer but not E is $(2)(48) = 96$, where 2 is the number of positions D can take and 48 is the number of selections of the other officer from the remaining people in the club except E. The number of selections when E serves as an officer but not D is also $(2)(48) = 96$. The number of selections when both D and E are not chosen is $_{48}P_2 = 2256$. Therefore, the total number of choices is $(2)(96) + 2256 = 2448$. This problem also has another short solution: Since D and E can only serve together in 2 ways, the answer is $2450 - 2 = 2448$.

Permutations that occur by arranging objects in a circle are called **circular permutations**. Two circular permutations are not considered different unless corresponding objects in the two arrangements are preceded or followed by a different object as we proceed in a clockwise direction. For example, if 4 people are playing bridge, we do not have a new permutation if they all move one position in a clockwise direction. By considering one person in a fixed position and arranging the other three in 3! ways, we find that there are 6 distinct arrangements for the bridge game.

Theorem 1.3: The number of permutations of n objects arranged in a circle is $(n-1)!$.

So far we have considered permutations of distinct objects. That is, all the objects were completely different or distinguishable. Obviously, if the letters b and c are both equal to x, then the 6 permutations of the letters a, b, and c become axx, axx, xax, xax, xxa, and xxa, of which only 3 are distinct. Therefore, with 3 letters, 2 being the same, we have $3!/2! = 3$ distinct permutations. With 4 different letters a, b, c, and d, we have 24 distinct permutations. If we let $a = b = x$ and $c = d = y$, we can list only the following distinct permutations: $xxyy$, $xyxy$, $yxxy$, $yyxx$, $xyyx$, and $yxyx$. Thus, we have $4!/(2!\,2!) = 6$ distinct permutations.

Theorem 1.4: The number of distinct permutations of n things of which n_1 are of one kind, n_2 of a second kind, $\ldots$, n_k of a kth kind is

$$\frac{n!}{n_1!n_2!\cdots n_k!}.$$

Example 1.17: In a college football training session, the defensive coordinator needs to have 10 players standing in a row. Among these 10 players, there are 1 freshman, 2 sophomores, 4 juniors, and 3 seniors. How many different ways can they be arranged in a row if only their class level will be distinguished?

***Solution*:** Directly using Theorem 1.4, we find that the total number of arrangements is

$$\frac{10!}{1!\ 2!\ 4!\ 3!} = 12{,}600.$$

Often we are concerned with the number of ways of partitioning a set of n objects into r subsets called **cells**. A partition has been achieved if the intersection of every possible pair of the r subsets is the empty set ϕ and if the union of all subsets gives the original set. The order of the elements within a cell is of no importance. Consider the set $\{a, e, i, o, u\}$. The possible partitions into two cells in which the first cell contains 4 elements and the second cell 1 element are

$$\{(a,e,i,o),(u)\}, \{(a,i,o,u),(e)\}, \{(e,i,o,u),(a)\}, \{(a,e,o,u),(i)\}, \{(a,e,i,u),(o)\}.$$

We see that there are 5 ways to partition a set of 4 elements into two subsets, or cells, containing 4 elements in the first cell and 1 element in the second.

The number of partitions for this illustration is denoted by the symbol

$$\binom{5}{4,1} = \frac{5!}{4!\ 1!} = 5,$$

where the top number represents the total number of elements and the bottom numbers represent the number of elements going into each cell. We state this more generally in Theorem 1.5.

Theorem 1.5: The number of ways of partitioning a set of n objects into r cells with n_1 elements in the first cell, n_2 elements in the second, and so forth, is

$$\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1! n_2! \cdots n_r!},$$

where $n_1 + n_2 + \cdots + n_r = n$.

Example 1.18: In how many ways can 7 graduate students be assigned to 1 triple and 2 double hotel rooms during a conference?

***Solution*:** The total number of possible partitions would be

$$\binom{7}{3, 2, 2} = \frac{7!}{3!\ 2!\ 2!} = 210.$$

In many problems, we are interested in the number of ways of selecting r objects from n without regard to order. These selections are called **combinations**. A combination is actually a partition with two cells, the one cell containing the r objects selected and the other cell containing the $(n-r)$ objects that are left. The number of such combinations, denoted by

$$\binom{n}{r, n-r}, \text{ is usually shortened to } \binom{n}{r},$$

since the number of elements in the second cell must be $n - r$.

Theorem 1.6: The number of combinations of n distinct objects taken r at a time is

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

Example 1.19: A young boy asks his mother to get 5 Game-Boy™ cartridges from his collection of 10 arcade and 5 sports game cartridges. How many ways are there that his mother can get 3 arcade and 2 sports game cartridges?

***Solution*:** The number of ways of selecting 3 cartridges from 10 is

$$\binom{10}{3} = \frac{10!}{3!\ (10-3)!} = 120.$$

The number of ways of selecting 2 cartridges from 5 is

$$\binom{5}{2} = \frac{5!}{2!\ 3!} = 10.$$

Hence for the total we have $(120)(10) = 1200$ ways.

Example 1.20: How many different letter arrangements can be made from the letters in the word *STATISTICS*?

***Solution*:** Using the same argument as in the discussion for Theorem 1.6, in this example we can actually apply Theorem 1.5 to obtain

$$\binom{10}{3, 3, 2, 1, 1} = \frac{10!}{3!\ 3!\ 2!\ 1!\ 1!} = 50{,}400.$$

Here we have 10 total letters, with 2 letters (S, T) appearing 3 times each, letter I appearing twice, and letters A and C appearing once each. Or this result can be obtained directly by using Theorem 1.4.

Exercises

1.17 Registrants at a large convention are offered 6 sightseeing tours on each of 3 days. In how many ways can a person arrange to go on a sightseeing tour planned by this convention?

1.18 In a medical study, patients are classified in 8 ways according to whether they have blood type AB^+, AB^-, A^+, A^-, B^+, B^-, O^+, or O^-, and also according to whether their blood pressure is low, normal, or high. Find the number of ways in which a patient can be classified.

1.19 Students at a private liberal arts college are classified as being freshmen, sophomores, juniors, or seniors, and also according to whether they are male or female. Find the total number of possible classifications for the students of that college.

1.20 A California study concluded that following 7 simple health rules can extend a man's life by 11 years on the average and a woman's life by 7 years. These 7 rules are as follows: no smoking, get regular exercise, use alcohol only in moderation, get 7 to 8 hours of sleep, maintain proper weight, eat breakfast, and do not eat between meals. In how many ways can a person adopt 5 of these rules to follow

(a) if the person presently violates all 7 rules?

(b) if the person never drinks and always eats breakfast?

1.21 A developer of a new subdivision offers a prospective home buyer a choice of 4 designs, 3 different heating systems, a garage or carport, and a patio or screened porch. How many different plans are available to this buyer?

1.22 A drug for the relief of asthma can be purchased from 5 different manufacturers in liquid, tablet, or capsule form, all of which come in regular and extra strength. How many different ways can a doctor prescribe the drug for a patient suffering from asthma?

1.23 In a fuel economy study, each of 3 race cars is tested using 5 different brands of gasoline at 7 test sites located in different regions of the country. If 2 drivers are used in the study, and test runs are made once under each distinct set of conditions, how many test runs are needed?

1.24 In how many different ways can a true-false test consisting of 9 questions be answered?

1.25 A witness to a hit-and-run accident told the police that the license number contained the letters RLH followed by 3 digits, the first of which was a 5. If the witness cannot recall the last 2 digits, but is certain that all 3 digits are different, find the maximum number of automobile registrations that the police may have to check.

1.26 (a) In how many ways can 6 people be lined up to get on a bus?

(b) If 3 specific persons, among 6, insist on following each other, how many ways are possible?

(c) If 2 specific persons, among 6, refuse to follow each other, how many ways are possible?

1.27 A contractor wishes to build 9 houses, each different in design. In how many ways can he place these houses on a street if 6 lots are on one side of the street and 3 lots are on the opposite side?

1.28 (a) How many distinct permutations can be made from the letters of the word *COLUMNS*?

(b) How many of these permutations start with the letter M?

1.29 In how many ways can 4 boys and 5 girls sit in a row if the boys and girls must alternate?

1.30 (a) How many three-digit numbers can be formed from the digits 0, 1, 2, 3, 4, 5, and 6 if each digit can be used only once?

(b) How many of these are odd numbers?

(c) How many are greater than 330?

1.31 In a regional spelling bee, the 8 finalists consist of 3 boys and 5 girls. Find the number of sample points in the sample space S for the number of possible orders at the conclusion of the contest for

(a) all 8 finalists;

(b) the first 3 positions.

1.32 Four married couples have bought 8 seats in the same row for a concert. In how many different ways can they be seated

(a) with no restrictions?

(b) if each couple is to sit together?

(c) if all the men sit together to the right of all the women?

1.33 Find the number of ways that 6 teachers can be assigned to 4 sections of an introductory psychology course if no teacher is assigned to more than one section.

1.34 Three lottery tickets for first, second, and third prizes are drawn from a group of 40 tickets. Find the number of sample points in S for awarding the 3 prizes if each contestant holds only 1 ticket.

1.35 In how many ways can 5 different trees be planted in a circle?

1.36 In how many ways can 3 oaks, 4 pines, and 2 maples be arranged along a property line if one does not distinguish among trees of the same kind?

1.37 How many ways are there that no two students will have the same birth date in a class of size 60?

1.6 Probability of an Event

Perhaps it was humankind's unquenchable thirst for gambling that led to the early development of probability theory. In an effort to increase their winnings, gamblers called upon mathematicians to provide optimum strategies for various games of chance. Some of the mathematicians providing these strategies were Pascal, Leibniz, Fermat, and James Bernoulli. As a result of this development of probability theory, statistical inference, with all its predictions and generalizations, has branched out far beyond games of chance to encompass many other fields associated with chance occurrences, such as politics, business, weather forecasting, and scientific research. For these predictions and generalizations to be reasonably accurate, an understanding of basic probability theory is essential.

What do we mean when we make the statement "John will probably win the tennis match," or "I have a fifty-fifty chance of getting an even number when a die is tossed," or "The university is not likely to win the football game tonight," or "Most of our graduating class will likely be married within 3 years"? In each case, we are expressing an outcome of which we are not certain, but owing to past information or from an understanding of the structure of the experiment, we have some degree of confidence in the validity of the statement.

Throughout the remainder of this chapter, we consider only those experiments for which the sample space contains a finite number of elements. The likelihood of the occurrence of an event resulting from such a statistical experiment is evaluated by means of a set of real numbers, called **weights** or **probabilities**, ranging from 0 to 1. To every point in the sample space we assign a probability such that the sum of all probabilities is 1. If we have reason to believe that a certain sample point is quite likely to occur when the experiment is conducted, the probability assigned should be close to 1. On the other hand, a probability closer to 0 is assigned to a sample point that is not likely to occur. In many experiments, such as tossing a fair coin or a balanced die, all the sample points have the same chance of occurring and are assigned equal probabilities. For points outside the sample space, that is, for simple events that cannot possibly occur, we assign a probability of 0.

To find the probability of an event A, we sum all the probabilities assigned to the sample points in A. This sum is called the **probability** of A and is denoted by $P(A)$.

Definition 1.8: The **probability** of an event A is the sum of the weights of all sample points in A. Therefore,

$$0 \le P(A) \le 1, \quad P(\phi) = 0, \quad \text{and} \quad P(S) = 1.$$

Furthermore, if $A_1, A_2, A_3, \ldots$ is a sequence of mutually exclusive events, then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots.$$

Example 1.21: A coin is tossed twice. What is the probability that at least 1 head occurs?

***Solution*:** The sample space for this experiment is

$$S = \{HH, HT, TH, TT\}.$$

If the coin is balanced, each of these outcomes is equally likely to occur. Therefore, we assign a probability of ω to each sample point. Then $4\omega = 1$, or $\omega = 1/4$. If A represents the event of at least 1 head occurring, then

$$A = \{HH, HT, TH\} \text{ and } P(A) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}.$$

Example 1.22: A die is loaded in such a way that an even number is twice as likely to occur as an odd number. If E is the event that a number less than 4 occurs on a single toss of the die, find $P(E)$.

***Solution*:** The sample space is $S = \{1, 2, 3, 4, 5, 6\}$. We assign a probability of w to each odd number and a probability of $2w$ to each even number. Since the sum of the probabilities must be 1, we have $9w = 1$, or $w = 1/9$. Hence, probabilities of 1/9 and 2/9 are assigned to each odd and even number, respectively. Therefore,

$$E = \{1, 2, 3\} \text{ and } P(E) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}.$$

If the sample space for an experiment contains N elements, all of which are equally likely to occur, we assign a probability equal to $1/N$ to each of the N points. The probability of any event A containing n of these N sample points is then the ratio of the number of elements in A to the number of elements in S.

Definition 1.9: If an experiment can result in any one of N different equally likely outcomes, and if exactly n of these outcomes correspond to event A, then the probability of event A is

$$P(A) = \frac{n}{N}.$$

Example 1.23: A statistics class for engineers consists of 25 industrial, 10 mechanical, 10 electrical, and 8 civil engineering students. If a person is randomly selected by the instructor to answer a question, find the probability that the student chosen is (a) an industrial engineering major and (b) a civil engineering or an electrical engineering major.

***Solution*:** Denote by I, M, E, and C the students majoring in industrial, mechanical, electrical, and civil engineering, respectively. The total number of students in the class is 53, all of whom are equally likely to be selected.

(a) Since 25 of the 53 students are majoring in industrial engineering, the probability of event I, selecting an industrial engineering major at random, is

$$P(I) = \frac{25}{53}.$$

(b) Since 18 of the 53 students are civil or electrical engineering majors, it follows that

$$P(C \cup E) = \frac{18}{53}.$$

1.7 Additive Rules

Often it is easiest to calculate the probability of some event from known probabilities of other events. This may well be true if the event in question can be represented as the union of two other events or as the complement of some event. Several important laws that frequently simplify the computation of probabilities follow. The first, called the **additive rule**, applies to unions of events.

Theorem 1.7: If A and B are two events, then

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

***Proof*:** Consider the Venn diagram in Figure 1.8. The $P(A \cup B)$ is the sum of the probabilities of the sample points in $A \cup B$. Now $P(A) + P(B)$ is the sum of all the probabilities in A plus the sum of all the probabilities in B. Therefore, we have added the probabilities in $(A \cap B)$ twice. Since these probabilities add up to $P(A \cap B)$, we must subtract this probability once to obtain the sum of the probabilities in $A \cup B$.

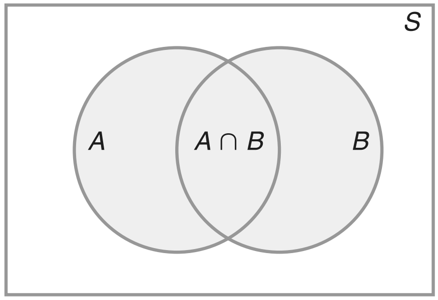

Figure 1.8: Additive rule of probability.

Corollary 1.1: If A and B are mutually exclusive, then

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

Corollary 1.1 is an immediate result of Theorem 1.7, since if A and B are mutually exclusive, $A \cap B = 0$ and then $P(A \cap B) = P(\phi) = 0$. In general, we can write Corollary 1.2.

Corollary 1.2: If $A_1,\ A_2, \ldots, A_n$ are mutually exclusive, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n).$$

A collection of events $\{A_1, A_2, \ldots, A_n\}$ of a sample space S is called a **partition** of S if $A_1, A_2, \ldots, A_n$ are mutually exclusive and $A_1 \cup A_2 \cup \cdots \cup A_n = S$. Thus, we have

Corollary 1.3: If $A_1,\ A_2, \ldots, A_n$ is a partition of sample space S, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n) = P(S) = 1.$$

As one might expect, Theorem 1.7 extends in an analogous fashion.

Theorem 1.8: For three events A, B, and C,

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\ &\quad - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C). \end{aligned}$$

Example 1.24: John is going to graduate from an industrial engineering department in a university by the end of the semester. After being interviewed at two companies he likes, he assesses that his probability of getting an offer from company A is 0.8, and his probability of getting an offer from company B is 0.6. If he believes that the probability that he will get offers from both companies is 0.5, what is the probability that he will get at least one offer from these two companies?

***Solution*:** Using the additive rule, we have

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.8 + 0.6 - 0.5 = 0.9.$$

Example 1.25: What is the probability of getting a total of 7 or 11 when a pair of fair dice is tossed?

***Solution*:** Let A be the event that 7 occurs and B the event that 11 comes up. Now, a total of 7 occurs for 6 of the 36 sample points, and a total of 11 occurs for only 2 of the sample points. Since all sample points are equally likely, we have $P(A) = 1/6$ and $P(B) = 1/18$. The events A and B are mutually exclusive, since a total of 7 and 11 cannot both occur on the same toss. Therefore,

$$P(A \cup B) = P(A) + P(B) = \frac{1}{6} + \frac{1}{18} = \frac{2}{9}.$$

This result could also have been obtained by counting the total number of points for the event $A \cup B$, namely 8, and writing

$$P(A \cup B) = \frac{n}{N} = \frac{8}{36} = \frac{2}{9}.$$

Theorem 1.7 and its three corollaries should help the reader gain more insight into probability and its interpretation. Corollaries 1.1 and 1.2 suggest the very intuitive result dealing with the probability of occurrence of at least one of a number of events, no two of which can occur simultaneously. The probability that at least one occurs is the sum of the probabilities of occurrence of the individual events. The third corollary simply states that the highest value of a probability (unity) is assigned to the entire sample space S.

Example 1.26: If the probabilities are, respectively, 0.09, 0.15, 0.21, and 0.23 that a person purchasing a new automobile will choose the color green, white, red, or blue, what is the probability that a given buyer will purchase a new automobile that comes in one of those colors?

***Solution*:** Let G, W, R, and B be the events that a buyer selects, respectively, a green, white, red, or blue automobile. Since these four events are mutually exclusive, the probability is

$$\begin{aligned} P(G \cup W \cup R \cup B) &= P(G) + P(W) + P(R) + P(B) \\ &= 0.09 + 0.15 + 0.21 + 0.23 = 0.68. \end{aligned}$$

Often it is more difficult to calculate the probability that an event occurs than it is to calculate the probability that the event does not occur. Should this be the case for some event A, we simply find $P(A')$ first and then, using Theorem 1.9, find $P(A)$ by subtraction.

Theorem 1.9: If A and A' are complementary events, then

$$P(A) + P(A') = 1.$$

***Proof*:** Since $A \cup A' = S$ and the sets A and A' are disjoint,

$$1 = P(S) = P(A \cup A') = P(A) + P(A').$$

Example 1.27: If the probabilities that an automobile mechanic will service 3, 4, 5, 6, 7, or 8 or more cars on any given workday are, respectively, 0.12, 0.19, 0.28, 0.24, 0.10, and 0.07, what is the probability that he will service at least 5 cars on his next day at work?

***Solution*:** Let E be the event that at least 5 cars are serviced. Now, $P(E) = 1 - P(E')$, where E' is the event that fewer than 5 cars are serviced. Since

$$P(E') = 0.12 + 0.19 = 0.31,$$

it follows from Theorem 1.9 that

$$P(E) = 1 - 0.31 = 0.69.$$

Example 1.28: Suppose the manufacturer's specifications for the length of a certain type of computer cable are 2000 ± 10 millimeters. In this industry, it is known that short cable is just as likely to be defective (not meeting specifications) as long cable. That is, the probability of randomly producing a cable with length exceeding 2010 millimeters is equal to the probability of producing a cable with length smaller than 1990 millimeters. The probability that the production procedure meets specifications is known to be 0.99.

(a) What is the probability that a cable selected randomly is too long?

(b) What is the probability that a randomly selected cable is longer than 1990 millimeters?

***Solution*:** Let M be the event that a cable meets specifications. Let S and L be the events that the cable is too short and too long, respectively. Then

(a) $P(M) = 0.99$ and $P(S) = P(L) = (1 - 0.99)/2 = 0.005$.

(b) Denoting by X the length of a randomly selected cable, we have

$$P(1990 \leq X \leq 2010) = P(M) = 0.99.$$

Since $P(X \geq 2010) = P(L) = 0.005$,

$$P(X \geq 1990) = P(M) + P(L) = 0.995.$$

This also can be solved by using Theorem 1.9:

$$P(X \geq 1990) + P(X < 1990) = 1.$$

Thus, $P(X \geq 1990) = 1 - P(S) = 1 - 0.005 = 0.995$.

Exercises

1.38 Suppose that in a college senior class of 500 students it is found that 210 smoke, 258 drink alcoholic beverages, 216 eat between meals, 122 smoke and drink alcoholic beverages, 83 eat between meals and drink alcoholic beverages, 97 smoke and eat between meals, and 52 engage in all three of these bad health practices. If a member of this senior class is selected at random, find the probability that the student

(a) smokes but does not drink alcoholic beverages;

(b) eats between meals and drinks alcoholic beverages but does not smoke;

(c) neither smokes nor eats between meals.

1.39 Find the errors in each of the following statements:

(a) The probabilities that an automobile salesperson will sell 0, 1, 2, or 3 cars on any given day in February are, respectively, 0.19, 0.38, 0.29, and 0.15.

(b) The probability that it will rain tomorrow is 0.40, and the probability that it will not rain tomorrow is 0.52.

(c) The probabilities that a printer will make 0, 1, 2, 3, or 4 or more mistakes in setting a document are, respectively, $0.19, 0.34, -0.25, 0.43$, and 0.29.

(d) On a single draw from a deck of playing cards, the probability of selecting a heart is $1/4$, the probability of selecting a black card is $1/2$, and the probability of selecting both a heart and a black card is $1/8$.

1.40 An automobile manufacturer is concerned about a possible recall of its best-selling four-door sedan. If there were a recall, there is a probability of 0.25 of a defect in the brake system, 0.18 of a defect in the transmission, 0.17 of a defect in the fuel system, and 0.40 of a defect in some other area.

(a) What is the probability that the defect is in the brakes or the fueling system if the probability of defects in both systems simultaneously is 0.15?

(b) What is the probability that there are no defects in either the brakes or the fueling system?

1.41 The probability that an American industry will locate in Shanghai, China, is 0.7, the probability that it will locate in Beijing, China, is 0.4, and the probability that it will locate in either Shanghai or Beijing or both is 0.8. What is the probability that the industry will locate

(a) in both cities?

(b) in neither city?

1.42 From past experience, a stockbroker believes that under present economic conditions a customer will invest in tax-free bonds with a probability of 0.6, will invest in mutual funds with a probability of 0.3, and will invest in both tax-free bonds and mutual funds with a probability of 0.15. At this time, find the probability that a customer will invest

(a) in either tax-free bonds or mutual funds;

(b) in neither tax-free bonds nor mutual funds.

1.43 A box contains 500 envelopes, of which 75 contain \$100 in cash, 150 contain \$25, and 275 contain \$10. An envelope may be purchased for \$25. What is the sample space for the different amounts of money? Assign probabilities to the sample points and then find the probability that the first envelope purchased contains less than \$100.

1.44 If 3 books are picked at random from a shelf containing 5 novels, 3 books of poems, and a dictionary, what is the probability that

(a) the dictionary is selected?

(b) 2 novels and 1 book of poems are selected?

1.45 In a high school graduating class of 100 students, 54 studied mathematics, 69 studied history, and 35 studied both mathematics and history. If one of these students is selected at random, find the probability that

(a) the student took mathematics or history;

(b) the student did not take either of these subjects;

(c) the student took history but not mathematics.

1.46 Dom's Pizza Company uses taste testing and statistical analysis of the data prior to marketing any new product. Consider a study involving three types of crusts (thin, thin with garlic and oregano, and thin with bits of cheese). Dom's is also studying three sauces (standard, a new sauce with more garlic, and a new sauce with fresh basil).

(a) How many combinations of crust and sauce are involved?

(b) What is the probability that a judge will get a plain thin crust with a standard sauce for his first taste test?

1.47 According to *Consumer Digest* (July/August 1996), the probable location of personal computers (PCs) in the home is as follows:

Adult bedroom:	0.03
Child bedroom:	0.15
Other bedroom:	0.14
Office or den:	0.40
Other rooms:	0.28

(a) What is the probability that a PC is in a bedroom?

(b) What is the probability that it is not in a bedroom?

(c) Suppose a household is selected at random from households with a PC; in what room would you expect to find a PC?

1.48 Interest centers around the life of an electronic component. Suppose it is known that the probability that the component survives for more than 6000 hours is 0.42. Suppose also that the probability that the component survives *no longer than* 4000 hours is 0.04.

(a) What is the probability that the life of the component is less than or equal to 6000 hours?

(b) What is the probability that the life is greater than 4000 hours?

1.49 Consider the situation of Exercise 1.48. Let A be the event that the component fails a particular test and B be the event that the component displays strain but does not actually fail. Event A occurs with probability 0.20, and event B occurs with probability 0.35.

(a) What is the probability that the component does not fail the test?

(b) What is the probability that the component works perfectly well (i.e., neither displays strain nor fails the test)?

(c) What is the probability that the component either fails or shows strain in the test?

1.50 Factory workers are constantly encouraged to practice zero tolerance when it comes to accidents in factories. Accidents can occur because the working environment or conditions themselves are unsafe. On the other hand, accidents can occur due to carelessness or so-called human error. In addition, the worker's shift, 7:00 A.M.–3:00 P.M. (day shift), 3:00 P.M.–11:00 P.M. (evening shift), or 11:00 P.M.–7:00 A.M. (graveyard shift), may be a factor. During the last year, 300 accidents have occurred. The percentages of the accidents for the condition combinations are as follows:

Shift	**Unsafe Conditions**	**Human Error**
Day	5%	32%
Evening	6%	25%
Graveyard	2%	30%

If an accident report is selected randomly from the 300 reports,

(a) what is the probability that the accident occurred on the graveyard shift?

(b) what is the probability that the accident occurred due to human error?

(c) what is the probability that the accident occurred due to unsafe conditions?

(d) what is the probability that the accident occurred on either the evening or the graveyard shift?

1.51 Consider the situation of Example 1.27 on page 30.

(a) What is the probability that no more than 4 cars will be serviced by the mechanic?

(b) What is the probability that he will service fewer than 8 cars?

(c) What is the probability that he will service either 3 or 4 cars?

1.52 Interest centers around the nature of an oven purchased at a particular department store. It can be either a gas or an electric oven. Consider the decisions made by six distinct customers.

(a) Suppose that the probability is 0.40 that at most two of these individuals purchase an electric oven. What is the probability that at least three purchase the electric oven?

(b) Suppose it is known that the probability that all six purchase the electric oven is 0.007 while 0.104 is the probability that all six purchase the gas oven. What is the probability that at least one of each type is purchased?

1.53 It is common in many industrial areas to use a filling machine to fill boxes full of product. This occurs in the food industry as well as other areas in which the product is used in the home (for example, detergent). These machines are not perfect, and indeed they may A, fill to specification, B, underfill, and C, overfill. Generally, the practice of underfilling is that which one hopes to avoid. Let $P(B) = 0.001$ while $P(A) = 0.990$.

(a) Give $P(C)$.

(b) What is the probability that the machine does not underfill?

(c) What is the probability that the machine either overfills or underfills?

1.54 Consider the situation of Exercise 1.53. Suppose 50,000 boxes of detergent are produced per week and suppose also that those underfilled are "sent back," with customers requesting reimbursement of the purchase price. Suppose also that the cost of production is known to be \$4.00 per box while the purchase price is \$4.50 per box.

(a) What is the weekly profit under the condition of no defective boxes?

(b) What is the loss in profit expected due to underfilling?

1.55 As the situation of Exercise 1.53 might suggest, statistical procedures are often used for control of quality (i.e., industrial quality control). At times, the *weight* of a product is an important variable to control. Specifications are given for the weight of a certain packaged product, and a package is rejected if it is either too light or too heavy. Historical data suggest that 0.95 is the probability that the product meets weight specifications whereas 0.002 is the probability that the product is too light. For each single packaged product, the manufacturer invests \$20.00 in production and the purchase price for the consumer is \$25.00.

(a) What is the probability that a package chosen randomly from the production line is too heavy?

(b) For each 10,000 packages sold, what profit is received by the manufacturer if all packages meet weight specification?

(c) Assuming that all defective packages are rejected and rendered worthless, how much is the profit reduced on 10,000 packages due to failure to meet weight specification?

1.56 Prove that

$$P(A' \cap B') = 1 + P(A \cap B) - P(A) - P(B).$$

1.8 Conditional Probability, Independence, and the Product Rule

One very important concept in probability theory is conditional probability. In some applications, the practitioner is interested in the probability structure under certain restrictions. For instance, in epidemiology, rather than studying the chance that a person from the general population has diabetes, it might be of more interest to know this probability for a distinct group such as Asian women in the age range of 35 to 50 or Hispanic men in the age range of 40 to 60. This type of probability is called a conditional probability.

Conditional Probability

The probability of an event B occurring when it is known that some event A has occurred is called a **conditional probability** and is denoted by $P(B|A)$. The symbol $P(B|A)$ is usually read "the probability that B occurs given that A occurs" or simply "the probability of B, given A."

Consider the event B of getting a perfect square when a die is tossed. The die is constructed so that the even numbers are twice as likely to occur as the odd numbers. Based on the sample space $S = \{1, 2, 3, 4, 5, 6\}$, with probabilities of 1/9 and 2/9 assigned, respectively, to the odd and even numbers, the probability of B occurring is 1/3. Now suppose that it is known that the toss of the die resulted in a number greater than 3. We are now dealing with a reduced sample space $A = \{4, 5, 6\}$, which is a subset of S. To find the probability that B occurs, relative to the space A, we must first assign new probabilities to the elements of A proportional to their original probabilities such that their sum is 1. Assigning a probability of w to the odd number in A and a probability of $2w$ to the two even numbers, we have $5w = 1$, or $w = 1/5$. Relative to the space A, we find that B contains the single element 4. Denoting this event by the symbol $B|A$, we write $B|A = \{4\}$, and hence

$$P(B|A) = \frac{2}{5}.$$

This example illustrates that events may have different probabilities when considered relative to different sample spaces.

We can also write

$$P(B|A) = \frac{2}{5} = \frac{2/9}{5/9} = \frac{P(A \cap B)}{P(A)},$$

where $P(A \cap B)$ and $P(A)$ are found from the original sample space S. In other words, a conditional probability relative to a subspace A of S may be calculated directly from the probabilities assigned to the elements of the original sample space S.

Definition 1.10: The conditional probability of B, given A, denoted by $P(B|A)$, is defined by

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad \text{provided} \quad P(A) > 0.$$

As an additional illustration, suppose that our sample space S is the population of adults in a small town who have completed the requirements for a college degree. We shall categorize them according to gender and employment status. The data are given in Table 1.3.

Table 1.3: Categorization of the Adults in a Small Town

	Employed	**Unemployed**	**Total**
Male	460	40	500
Female	140	260	400
Total	600	300	900

One of these individuals is to be selected at random for a tour throughout the country to publicize the advantages of establishing new industries in the town. We shall be concerned with the following events:

M: a man is chosen,

E: the one chosen is employed.

Using the reduced sample space E, we find that

$$P(M|E) = \frac{460}{600} = \frac{23}{30}.$$

Let $n(A)$ denote the number of elements in any set A. Using this notation, since each adult has an equal chance of being selected, we can write

$$P(M|E) = \frac{n(E \cap M)}{n(E)} = \frac{n(E \cap M)/n(S)}{n(E)/n(S)} = \frac{P(E \cap M)}{P(E)},$$

where $P(E \cap M)$ and $P(E)$ are found from the original sample space S. To verify this result, note that

$$P(E) = \frac{600}{900} = \frac{2}{3} \quad \text{and} \quad P(E \cap M) = \frac{460}{900} = \frac{23}{45}.$$

Hence,

$$P(M|E) = \frac{23/45}{2/3} = \frac{23}{30},$$

as before.

Example 1.29: The probability that a regularly scheduled flight departs on time is $P(D) = 0.83$; the probability that it arrives on time is $P(A) = 0.82$; and the probability that it departs and arrives on time is $P(D \cap A) = 0.78$. Find the probability that a plane (a) arrives on time, given that it departed on time, and (b) departed on time, given that it has arrived on time.

***Solution*:** Using Definition 1.10, we have the following.

(a) The probability that a plane arrives on time, given that it departed on time, is

$$P(A|D) = \frac{P(D \cap A)}{P(D)} = \frac{0.78}{0.83} = 0.94.$$

(b) The probability that a plane departed on time, given that it has arrived on time, is

$$P(D|A) = \frac{P(D \cap A)}{P(A)} = \frac{0.78}{0.82} = 0.95.$$

The notion of conditional probability provides the capability of reevaluating the idea of probability of an event in light of additional information, that is, when it is known that another event has occurred. The probability $P(A|B)$ is an updating of $P(A)$ based on the knowledge that event B has occurred. In Example 1.29, it is important to know the probability that the flight arrives on time. One is given the information that the flight did not depart on time. Armed with this additional information, one can calculate the more pertinent probability $P(A|D')$, that is, the probability that it arrives on time, given that it did not depart on time. In many situations, the conclusions drawn from observing the more important conditional probability change the picture entirely. In this example, the computation of $P(A|D')$ is

$$P(A|D') = \frac{P(A \cap D')}{P(D')} = \frac{0.82 - 0.78}{0.17} = 0.24.$$

As a result, the probability of an on-time arrival is diminished severely in the presence of the additional information.

Example 1.30: The concept of conditional probability has countless uses in both industrial and biomedical applications. Consider an industrial process in the textile industry in which strips of a particular type of cloth are being produced. These strips can be defective in two ways, length and nature of texture. For the case of the latter, the process of identification is very complicated. It is known from historical information on the process that 10% of strips fail the length test, 5% fail the texture test, and only 0.8% fail both tests. If a strip is selected randomly from the process and a quick measurement identifies it as failing the length test, what is the probability that it is texture defective?

***Solution*:** Consider the events

$$L\text{: length defective,} \qquad T\text{: texture defective.}$$

Given that the strip is length defective, the probability that this strip is texture defective is given by

$$P(T|L) = \frac{P(T \cap L)}{P(L)} = \frac{0.008}{0.1} = 0.08.$$

Thus, knowing the conditional probability provides considerably more information than merely knowing $P(T)$.

Independent Events

In the die-tossing experiment discussed on page 33, we note that $P(B|A) = 2/5$ whereas $P(B) = 1/3$. That is, $P(B|A) \neq P(B)$, indicating that B depends on A. Now consider an experiment in which 2 cards are drawn in succession from an ordinary deck, with replacement. The events are defined as

A: the first card is an ace,

B: the second card is a spade.

Since the first card is replaced, our sample space for both the first and the second draw consists of 52 cards, containing 4 aces and 13 spades. Hence,

$$P(B|A) = \frac{13}{52} = \frac{1}{4} \quad \text{and} \quad P(B) = \frac{13}{52} = \frac{1}{4}.$$

That is, $P(B|A) = P(B)$. When this is true, the events A and B are said to be **independent**.

Although conditional probability allows for an alteration of the probability of an event in the light of additional material, it also enables us to understand better the very important concept of **independence** or, in the present context, independent events. In the airport illustration in Example 1.29, $P(A|D)$ differs from $P(A)$. This suggests that the occurrence of D influenced A, and this is certainly expected in this illustration. However, consider the situation where we have events A and B and

$$P(A|B) = P(A).$$

In other words, the occurrence of B had no impact on the chance of occurrence of A. Here the occurrence of A is independent of the occurrence of B. The importance of the concept of independence cannot be overemphasized. It plays a vital role in material in virtually all chapters in this book and in all areas of applied statistics.

Definition 1.11: Two events A and B are **independent** if and only if

$$P(B|A) = P(B) \quad \text{or} \quad P(A|B) = P(A),$$

assuming the existences of the conditional probabilities. Otherwise, A and B are **dependent**.

The condition $P(B|A) = P(B)$ implies that $P(A|B) = P(A)$, and conversely. For the card-drawing experiments, where we showed that $P(B|A) = P(B) = 1/4$, we also can see that $P(A|B) = P(A) = 1/13$.

The Product Rule, or the Multiplicative Rule

Multiplying the formula in Definition 1.10 by $P(A)$, we obtain the following important **multiplicative rule** (or **product rule**), which enables us to calculate the probability that two events will both occur.

Theorem 1.10: If in an experiment the events A and B can both occur, then

$$P(A \cap B) = P(A)P(B|A), \text{ provided } P(A) > 0.$$

Thus, the probability that both A and B occur is equal to the probability that A occurs multiplied by the conditional probability that B occurs, given that A occurs. Since the events $A \cap B$ and $B \cap A$ are equivalent, it follows from Theorem 1.10 that we can also write

$$P(A \cap B) = P(B \cap A) = P(B)P(A|B).$$

In other words, it does not matter which event is referred to as A and which event is referred to as B.

Example 1.31: Suppose that we have a fuse box containing 20 fuses, of which 5 are defective. If 2 fuses are selected at random and removed from the box in succession without replacing the first, what is the probability that both fuses are defective?

***Solution*:** We shall let A be the event that the first fuse is defective and B the event that the second fuse is defective; then we interpret $A \cap B$ as the event that A occurs and then B occurs after A has occurred. The probability of first removing a defective fuse is $1/4$; then the probability of removing a second defective fuse from the remaining 4 is $4/19$. Hence,

$$P(A \cap B) = \left(\frac{1}{4}\right)\left(\frac{4}{19}\right) = \frac{1}{19}.$$

If, in Example 1.31, the first fuse is replaced and the fuses thoroughly rearranged before the second is removed, then the probability of a defective fuse on the second selection is still $1/4$; that is, $P(B|A) = P(B)$ and the events A and B are independent. When this is true, we can substitute $P(B)$ for $P(B|A)$ in Theorem 1.10 to obtain the following special multiplicative rule.

Theorem 1.11: Two events A and B are independent if and only if

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

Therefore, to obtain the probability that two independent events will both occur, we simply find the product of their individual probabilities.

Example 1.32: A small town has one fire engine and one ambulance available for emergencies. The probability that the fire engine is available when needed is 0.98, and the probability that the ambulance is available when called is 0.92. In the event of an injury resulting from a burning building, find the probability that both the ambulance and the fire engine will be available, assuming they operate independently.

***Solution*:** Let A and B represent the respective events that the fire engine and the ambulance are available. Then

$$P(A \cap B) = P(A)P(B) = (0.98)(0.92) = 0.9016.$$

Example 1.33: An electrical system consists of four components as illustrated in Figure 1.9. The system works if components A and B work and either of the components C or D works. The reliability (probability of working) of each component is also shown in Figure 1.9. Find the probability that (a) the entire system works and (b) component C does not work, given that the entire system works. Assume that the four components work independently.

***Solution*:** In this configuration of the system, A, B, and the subsystem C and D constitute a serial circuit system, whereas the subsystem C and D itself is a parallel circuit system.

(a) Clearly the probability that the entire system works can be calculated as follows:

$$\begin{aligned} P[A \cap B \cap (C \cup D)] &= P(A)P(B)P(C \cup D) = P(A)P(B)[1 - P(C' \cap D')] \\ &= P(A)P(B)[1 - P(C')P(D')] \\ &= (0.9)(0.9)[1 - (1 - 0.8)(1 - 0.8)] = 0.7776. \end{aligned}$$

The equalities above hold because of the independence among the four components.

(b) To calculate the conditional probability in this case, notice that

$$\begin{aligned} P &= \frac{P(\text{the system works but } C \text{ does not work})}{P(\text{the system works})} \\ &= \frac{P(A \cap B \cap C' \cap D)}{P(\text{the system works})} = \frac{(0.9)(0.9)(1 - 0.8)(0.8)}{0.7776} = 0.1667. \end{aligned}$$

The multiplicative rule can be extended to more than two-event situations.

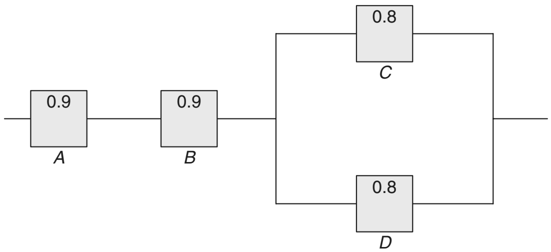

Figure 1.9: An electrical system for Example 1.33.

Theorem 1.12: If, in an experiment, the events $A_1, A_2, \ldots, A_k$ can occur, then

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_k|A_1 \cap A_2 \cap \cdots \cap A_{k-1}).$$

If the events $A_1, A_2, \ldots, A_k$ are independent, then

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2) \cdots P(A_k).$$

The property of independence stated in Theorem 1.11 can be extended to deal with more than two events. Consider, for example, the case of three events A, B, and C. It is not sufficient to only have that $P(A \cap B \cap C) = P(A)P(B)P(C)$ as a definition of independence among the three. Suppose $A = B$ and $C = \phi$, the null set. Although $A \cap B \cap C = \phi$, which results in $P(A \cap B \cap C) = 0 = P(A)P(B)P(C)$, events A and B are not independent. Hence, we have the following definition.

Definition 1.12: A collection of events $\mathcal{A} = \{A_1, \ldots, A_n\}$ are mutually independent if for any subset of $\mathcal{A}$, $A_{i_1}, \ldots, A_{i_k}$, for $k \leq n$, we have

$$P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k}).$$

Exercises

1.57 If R is the event that a convict committed armed robbery and D is the event that the convict sold drugs, state in words what probabilities are expressed by

(a) $P(R|D)$;

(b) $P(D'|R)$;

(c) $P(R'|D')$.

1.58 In an experiment to study the relationship of hypertension and smoking habits, the following data are collected for 180 individuals:

	Nonsmokers	Moderate Smokers	Heavy Smokers
H	21	36	30
NH	48	26	19

where H and NH in the table stand for *Hypertension* and *Nonhypertension*, respectively. If one of these individuals is selected at random, find the probability that the person is

(a) experiencing hypertension, given that the person is a heavy smoker;

(b) a nonsmoker, given that the person is experiencing no hypertension.

1.59 In *USA Today* (Sept. 5, 1996), the results of a survey involving the use of sleepwear while traveling were listed as follows:

	Male	**Female**	**Total**
Underwear	0.220	0.024	0.244
Nightgown	0.002	0.180	0.182
Nothing	0.160	0.018	0.178
Pajamas	0.102	0.073	0.175
T-shirt	0.046	0.088	0.134
Other	0.084	0.003	0.087

(a) What is the probability that a traveler is a female who sleeps in the nude?

(b) What is the probability that a traveler is male?

(c) Assuming the traveler is male, what is the probability that he sleeps in pajamas?

(d) What is the probability that a traveler is male if the traveler sleeps in pajamas or a T-shirt?

1.60 A manufacturer of a flu vaccine is concerned about the quality of its flu serum. Batches of serum are processed by three different departments having rejection rates of 0.10, 0.08, and 0.12, respectively. The inspections by the three departments are sequential and independent.

(a) What is the probability that a batch of serum survives the first departmental inspection but is rejected by the second department?

(b) What is the probability that a batch of serum is rejected by the third department?

1.61 The probability that a vehicle entering the Luray Caverns has Canadian license plates is 0.12; the probability that it is a camper is 0.28; and the probability that it is a camper with Canadian license plates is 0.09. What is the probability that

(a) a camper entering the Luray Caverns has Canadian license plates?

(b) a vehicle with Canadian license plates entering the Luray Caverns is a camper?

(c) a vehicle entering the Luray Caverns does not have Canadian plates or is not a camper?

1.62 For married couples living in a certain suburb, the probability that the husband will vote on a bond referendum is 0.21, the probability that the wife will vote on the referendum is 0.28, and the probability that both the husband and the wife will vote is 0.15. What is the probability that

(a) at least one member of a married couple will vote?

(b) a wife will vote, given that her husband will vote?

(c) a husband will vote, given that his wife will not vote?

1.63 The probability that a doctor correctly diagnoses a particular illness is 0.7. Given that the doctor makes an incorrect diagnosis, the probability that the patient files a lawsuit is 0.9. What is the probability that the doctor makes an incorrect diagnosis and the patient sues?

1.64 The probability that an automobile being filled with gasoline also needs an oil change is 0.25; the probability that it needs a new oil filter is 0.40; and the probability that both the oil and the filter need changing is 0.14.

(a) If the oil has to be changed, what is the probability that a new oil filter is needed?

(b) If a new oil filter is needed, what is the probability that the oil has to be changed?

1.65 In 1970, 11% of Americans completed four years of college; 43% of them were women. In 1990, 22% of Americans completed four years of college; 53% of them were women (*Time*, Jan. 19, 1996).

(a) Given that a person completed four years of college in 1970, what is the probability that the person was a woman?

(b) What is the probability that a woman finished four years of college in 1990?

(c) What is the probability that a man had not finished college in 1990?

1.66 Before the distribution of certain statistical software, every fourth compact disk (CD) is tested for accuracy. The testing process consists of running four independent programs and checking the results. The failure rates for the four testing programs are, respectively, 0.01, 0.03, 0.02, and 0.01.

(a) What is the probability that a CD was tested and failed any test?

(b) Given that a CD was tested, what is the probability that it failed program 2 or 3?

(c) In a sample of 100, how many CDs would you expect to be rejected?

(d) Given that a CD was defective, what is the probability that it was tested?

1.67 A town has two fire engines operating independently. The probability that a specific engine is available when needed is 0.96.

(a) What is the probability that neither is available when needed?

(b) What is the probability that a fire engine is available when needed?

1.68 Pollution of the rivers in the United States has been a problem for many years. Consider the following events:

A: the river is polluted,

B: a sample of water tested detects pollution,

C: fishing is permitted.

Assume $P(A) = 0.3$, $P(B|A) = 0.75$, $P(B|A') = 0.20$, $P(C|A \cap B) = 0.20$, $P(C|A' \cap B) = 0.15$, $P(C|A \cap B') = 0.80$, and $P(C|A' \cap B') = 0.90$.

(a) Find $P(A \cap B \cap C)$.

(b) Find $P(B' \cap C)$.

(c) Find $P(C)$.

(d) Find the probability that the river is polluted, given that fishing is permitted and the sample tested did not detect pollution.

1.69 A circuit system is given in Figure 1.10. Assume the components fail independently.

(a) What is the probability that the entire system works?

(b) Given that the system works, what is the probability that component A is not working?

1.70 Suppose the diagram of an electrical system is as given in Figure 1.11. What is the probability that the system works? Assume the components fail independently.

1.71 In the situation of Exercise 1.69, it is known that the system does not work. What is the probability that component A also does not work?

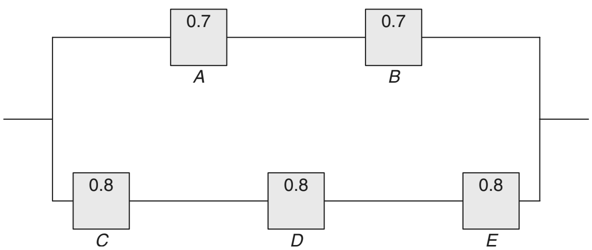

Figure 1.10: Diagram for Exercise 1.69.

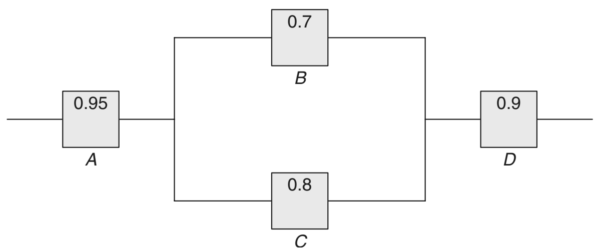

Figure 1.11: Diagram for Exercise 1.70.

1.9 Bayes' Rule

Bayesian statistics is a collection of tools that is used in a special form of statistical inference which applies in the analysis of experimental data in many practical situations in science and engineering. Bayes' rule is one of the most important rules in probability theory.

Total Probability

Let us now return to the illustration of Section 1.8, where an individual is being selected at random from the adults of a small town to tour the country and publicize the advantages of establishing new industries in the town. Suppose that we are now given the additional information that 36 of those employed and 12 of those unemployed are members of the Rotary Club. We wish to find the probability of the event A that the individual selected is a member of the Rotary Club. Referring to Figure 1.12, we can write A as the union of the two mutually exclusive events $E \cap A$ and $E' \cap A$. Hence, $A = (E \cap A) \cup (E' \cap A)$, and by Corollary 1.1 of Theorem 1.7, and then Theorem 1.10, we can write

$$\begin{aligned} P(A) &= P[(E \cap A) \cup (E' \cap A)] = P(E \cap A) + P(E' \cap A) \\ &= P(E)P(A|E) + P(E')P(A|E'). \end{aligned}$$

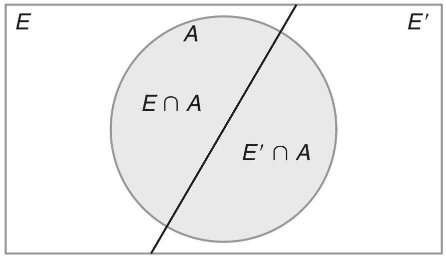

Figure 1.12: Venn diagram for the events A, E, and E'.

The data of Section 1.8, together with the additional data given above for the set A, enable us to compute

$$P(E) = \frac{600}{900} = \frac{2}{3}, \quad P(A|E) = \frac{36}{600} = \frac{3}{50},$$

and

$$P(E') = \frac{1}{3}, \quad P(A|E') = \frac{12}{300} = \frac{1}{25}.$$

If we display these probabilities by means of the tree diagram of Figure 1.13, where the first branch yields the probability $P(E)P(A|E)$ and the second branch yields the probability $P(E')P(A|E')$, it follows that

$$P(A) = \left(\frac{2}{3}\right)\left(\frac{3}{50}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{25}\right) = \frac{4}{75}.$$

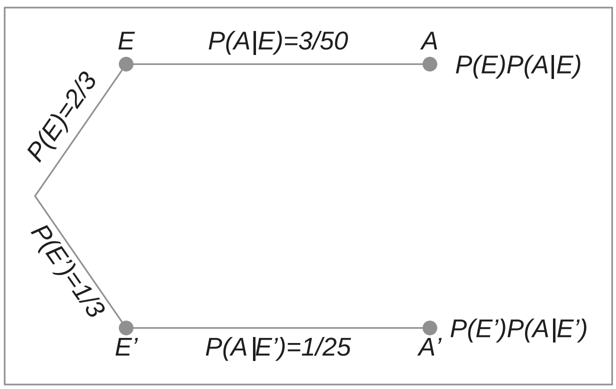

Figure 1.13: Tree diagram for the data on page 34, using additional information given above.

A generalization of the foregoing illustration to the case where the sample space is partitioned into k subsets is covered by the following theorem, sometimes called the **theorem of total probability** or the **rule of elimination**.

Theorem 1.13: If the events $B_1, B_2, \ldots, B_k$ constitute a partition of the sample space S such that $P(B_i) \neq 0$ for $i = 1, 2, \ldots, k$, then for any event A of S,

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i)P(A|B_i).$$

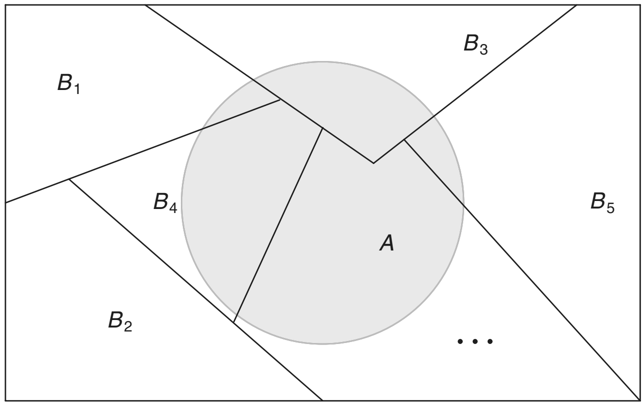

Figure 1.14: Partitioning the sample space S.

***Proof*:** Consider the Venn diagram of Figure 1.14. The event A is seen to be the union of the mutually exclusive events

$$B_1 \cap A, \; B_2 \cap A, \; \ldots, \; B_k \cap A;$$

that is,

$$A = (B_1 \cap A) \cup (B_2 \cap A) \cup \cdots \cup (B_k \cap A).$$

Using Corollary 1.2 of Theorem 1.7 and Theorem 1.10, we have

$$\begin{aligned} P(A) &= P[(B_1 \cap A) \cup (B_2 \cap A) \cup \cdots \cup (B_k \cap A)] \\ &= P(B_1 \cap A) + P(B_2 \cap A) + \cdots + P(B_k \cap A) \\ &= \sum_{i=1}^{k} P(B_i \cap A) \\ &= \sum_{i=1}^{k} P(B_i)P(A|B_i). \end{aligned}$$

Example 1.34: In a certain assembly plant, three machines, B_1, B_2, and B_3, make 30%, 45%, and 25%, respectively, of the products. It is known from past experience that 2%, 3%, and 2% of the products made by each machine, respectively, are defective. Now, suppose that a finished product is randomly selected. What is the probability that it is defective?

***Solution*:** Consider the following events:

A: the product is defective,

B_1: the product is made by machine B_1,

B_2: the product is made by machine B_2,

B_3: the product is made by machine B_3.

Applying the rule of elimination, we can write

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3).$$

Referring to the tree diagram of Figure 1.15, we find that the three branches give the probabilities

$$P(B_1)P(A|B_1) = (0.3)(0.02) = 0.006,$$
$$P(B_2)P(A|B_2) = (0.45)(0.03) = 0.0135,$$
$$P(B_3)P(A|B_3) = (0.25)(0.02) = 0.005,$$

and hence

$$P(A) = 0.006 + 0.0135 + 0.005 = 0.0245.$$

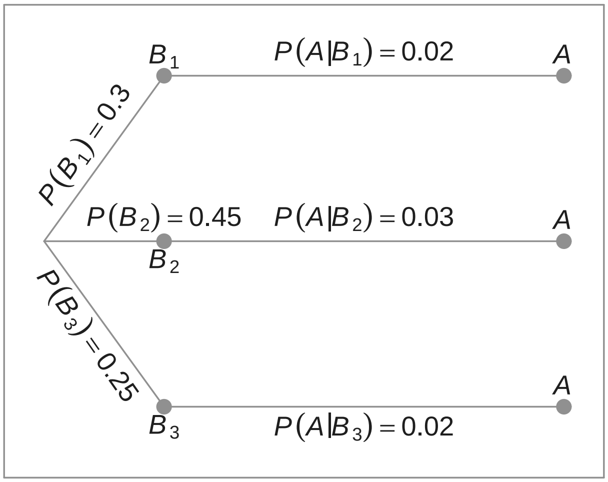

Figure 1.15: Tree diagram for Example 1.34.

Bayes' Rule

Instead of asking for $P(A)$ in Example 1.34, by the rule of elimination, suppose that we now consider the problem of finding the conditional probability $P(B_i|A)$. In other words, suppose that a product was randomly selected and it is defective. What is the probability that this product was made by machine B_i? Questions of this type can be answered by using the following theorem, called **Bayes' rule**:

Theorem 1.14: **(Bayes' Rule)** If the events $B_1, B_2, \ldots, B_k$ constitute a partition of the sample space S such that $P(B_i) \neq 0$ for $i = 1, 2, \ldots, k$, then for any event A in S such that $P(A) \neq 0$,

$$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)} \quad \text{for } r = 1, 2, \ldots, k.$$

***Proof*:** By the definition of conditional probability,

$$P(B_r|A) = \frac{P(B_r \cap A)}{P(A)},$$

and then using Theorem 1.13 in the denominator, we have

$$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)},$$

which completes the proof.

Example 1.35: With reference to Example 1.34, if a product was chosen randomly and found to be defective, what is the probability that it was made by machine B_3?

***Solution*:** Using Bayes' rule to write

$$P(B_3|A) = \frac{P(B_3)P(A|B_3)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)},$$

and then substituting the probabilities calculated in Example 1.34, we have

$$P(B_3|A) = \frac{0.005}{0.006 + 0.0135 + 0.005} = \frac{0.005}{0.0245} = \frac{10}{49}.$$

In view of the fact that a defective product was selected, this result suggests that it probably was not made by machine B_3.

Example 1.36: A manufacturing firm employs three analytical plans for the design and development of a particular product. For cost reasons, all three are used at varying times. In fact, plans 1, 2, and 3 are used for 30%, 20%, and 50% of the products, respectively. The defect rate is different for the three procedures as follows:

$$P(D|P_1) = 0.01, \qquad P(D|P_2) = 0.03, \qquad P(D|P_3) = 0.02,$$

where $P(D|P_j)$ is the probability of a defective product, given plan j. If a random product was observed and found to be defective, which plan was most likely used and thus responsible?

***Solution*:** From the statement of the problem

$$P(P_1) = 0.30, \quad P(P_2) = 0.20, \quad \text{and} \quad P(P_3) = 0.50,$$

we must find $P(P_j|D)$ for $j = 1, 2, 3$. Bayes' rule (Theorem 1.14) shows

$$P(P_1|D) = \frac{P(P_1)P(D|P_1)}{P(P_1)P(D|P_1) + P(P_2)P(D|P_2) + P(P_3)P(D|P_3)}$$
$$= \frac{(0.30)(0.01)}{(0.3)(0.01) + (0.20)(0.03) + (0.50)(0.02)} = \frac{0.003}{0.019} = 0.158.$$

Similarly,

$$P(P_2|D) = \frac{(0.03)(0.20)}{0.019} = 0.316 \text{ and } P(P_3|D) = \frac{(0.02)(0.50)}{0.019} = 0.526.$$

The conditional probability of a defect given plan 3 is the largest of the three; thus a defective for a random product is most likely the result of the use of plan 3. ∎

Exercises

1.72 Police plan to enforce speed limits by using radar traps at four different locations within the city limits. The radar traps at each of the locations L_1, L_2, L_3, and L_4 will be operated 40%, 30%, 20%, and 30% of the time. If a person who is speeding on her way to work has probabilities of 0.2, 0.1, 0.5, and 0.2, respectively, of passing through these locations, what is the probability that she will receive a speeding ticket?

1.73 In a certain region of the country it is known from past experience that the probability of selecting an adult over 40 years of age with cancer is 0.05. If the probability of a doctor correctly diagnosing a person with cancer as having the disease is 0.78 and the probability of incorrectly diagnosing a person without cancer as having the disease is 0.06, what is the probability that an adult over 40 years of age is diagnosed as having cancer?

1.74 If the person in Exercise 1.72 received a speeding ticket on her way to work, what is the probability that she passed through the radar trap located at L_2?

1.75 Referring to Exercise 1.73, what is the probability that a person diagnosed as having cancer actually has the disease?

1.76 A regional telephone company operates three identical relay stations at different locations. During a one-year period, the number of malfunctions reported by each station and the causes are shown below.

Station	A	B	C
Problems with electricity supplied	2	1	1
Computer malfunction	4	3	2
Malfunctioning electrical equipment	5	4	2
Caused by other human errors	7	7	5

Suppose that a malfunction was reported and it was found to be caused by other human errors. What is the probability that it came from station C?

1.77 Suppose that the four inspectors at a film factory are supposed to stamp the expiration date on each package of film at the end of the assembly line. John, who stamps 20% of the packages, fails to stamp the expiration date once in every 200 packages; Tom, who stamps 60% of the packages, fails to stamp the expiration date once in every 100 packages; Jeff, who stamps 15% of the packages, fails to stamp the expiration date once in every 90 packages; and Pat, who stamps 5% of the packages, fails to stamp the expiration date once in every 200 packages. If a customer complains that her package of film does not show the expiration date, what is the probability that it was inspected by John?

1.78 Denote by A, B, and C the events that a grand prize is behind doors A, B, and C, respectively. Suppose you randomly picked a door, say A. The game host opened a door, say B, and showed there was no prize behind it. Now the host offers you the option of either staying at the door that you picked (A) or switching to the remaining unopened door (C). Use probability to explain whether you should switch or not.

1.79 A paint-store chain produces and sells latex and semigloss paint. Based on long-range sales, the probability that a customer will purchase latex paint is 0.75. Of those that purchase latex paint, 60% also purchase rollers. But only 30% of semigloss paint buyers purchase rollers. A randomly selected buyer purchases a roller and a can of paint. What is the probability that the paint is latex?

Review Exercises

1.80 A truth serum has the property that 90% of the guilty suspects are properly judged while, of course, 10% of the guilty suspects are improperly found innocent. On the other hand, innocent suspects are misjudged 1% of the time. If the suspect was selected from a group of suspects of which only 5% have ever committed a crime, and the serum indicates that he is guilty, what is the probability that he is innocent?

1.81 An allergist claims that 50% of the patients she tests are allergic to some type of weed. What is the probability that

(a) exactly 3 of her next 4 patients are allergic to weeds?

(b) none of her next 4 patients is allergic to weeds?

1.82 By comparing appropriate regions of Venn diagrams, verify that

(a) $(A \cap B) \cup (A \cap B') = A$;

(b) $A' \cap (B' \cup C) = (A' \cap B') \cup (A' \cap C)$.

1.83 The probabilities that a service station will pump gas into 0, 1, 2, 3, 4, or 5 or more cars during a certain 30-minute period are 0.03, 0.18, 0.24, 0.28, 0.10, and 0.17, respectively. Find the probability that in this 30-minute period

(a) more than 2 cars receive gas;

(b) at most 4 cars receive gas;

(c) 4 or more cars receive gas.

1.84 A large industrial firm uses three local motels to provide overnight accommodations for its clients. From past experience it is known that 20% of the clients are assigned rooms at the Ramada Inn, 50% at the Sheraton, and 30% at the Lakeview Motor Lodge. If the plumbing is faulty in 5% of the rooms at the Ramada Inn, in 4% of the rooms at the Sheraton, and in 8% of the rooms at the Lakeview Motor Lodge, what is the probability that

(a) a client will be assigned a room with faulty plumbing?

(b) a person with a room having faulty plumbing was assigned accommodations at the Lakeview Motor Lodge?

1.85 The probability that a patient recovers from a delicate heart operation is 0.8. What is the probability that

(a) exactly 2 of the next 3 patients who have this operation survive?

(b) all of the next 3 patients who have this operation survive?

1.86 In a certain federal prison, it is known that 2/3 of the inmates are under 25 years of age. It is also known that 3/5 of the inmates are male and that 5/8 of the inmates are female or 25 years of age or older. What is the probability that a prisoner selected at random from this prison is female and at least 25 years old?

1.87 A shipment of 12 television sets contains 3 defective sets. In how many ways can a hotel purchase 5 of these sets and receive at least 2 of the defective sets?

1.88 A certain federal agency employs three consulting firms (A, B, and C) with probabilities 0.40, 0.35, and 0.25, respectively. From past experience it is known that the probabilities of cost overruns for the firms are 0.05, 0.03, and 0.15, respectively. Suppose a cost overrun is experienced by the agency.

(a) What is the probability that the consulting firm involved is company C?

(b) What is the probability that it is company A?

1.89 A manufacturer is studying the effects of cooking temperature, cooking time, and type of cooking oil on making potato chips. Three different temperatures, 4 different cooking times, and 3 different oils are to be used.

(a) What is the total number of combinations to be studied?

(b) How many combinations will be used for each type of oil?

(c) Discuss why permutations are not an issue in this exercise.

1.90 Consider the situation in Exercise 1.89, and suppose that the manufacturer can try only two combinations in a day.

(a) What is the probability that any given set of two runs is chosen?

(b) What is the probability that the highest temperature is used in either of these two combinations?

1.91 A certain form of cancer is known to be found in women over 60 with probability 0.07. A blood test exists for the detection of the disease, but the test is not infallible. In fact, it is known that 10% of the time the test gives a false negative (i.e., the test incorrectly gives a negative result) and 5% of the time the test gives a false positive (i.e., incorrectly gives a positive result). If a woman over 60 is known to have taken the test and received a favorable (i.e., negative) result, what is the probability that she has the disease?

1.92 A producer of a certain type of electronic component ships to suppliers in lots of twenty. Suppose that 60% of all such lots contain no defective components, 30% contain one defective component, and 10% contain two defective components. A lot is picked, two components from the lot are randomly selected and tested, and neither is defective.

(a) What is the probability that zero defective components exist in the lot?

(b) What is the probability that one defective exists in the lot?

(c) What is the probability that two defectives exist in the lot?

1.93 A construction company employs two sales engineers. Engineer 1 does the work of estimating cost for 70% of jobs bid by the company. Engineer 2 does the work for 30% of jobs bid by the company. It is known that the error rate for engineer 1 is such that 0.02 is the probability of an error when he does the work, whereas the probability of an error in the work of engineer 2 is 0.04. Suppose a bid arrives and a serious error occurs in estimating cost. Which engineer would you guess did the work? Explain and show all work.

1.94 In the field of quality control, the science of statistics is often used to determine if a process is "out of control." Suppose the process is, indeed, out of control and 20% of items produced are defective.

(a) If three items arrive off the process line in succession, what is the probability that all three are defective?

(b) If four items arrive in succession, what is the probability that three are defective?

1.95 An industrial plant is conducting a study to determine how quickly injured workers are back on the job following injury. Records show that 10% of all injured workers are admitted to the hospital for treatment and 15% are back on the job the next day. In addition, studies show that 2% are both admitted for hospital treatment and back on the job the next day. If a worker is injured, what is the probability that the worker will either be admitted to a hospital or be back on the job the next day or both?

1.96 A firm is accustomed to training operators who do certain tasks on a production line. Those operators who attend the training course are known to be able to meet their production quotas 90% of the time. New operators who do not take the training course only meet their quotas 65% of the time. Fifty percent of new operators attend the course. Given that a new operator meets her production quota, what is the probability that she attended the program?

1.97 During bad economic times, industrial workers are dismissed and are often replaced by machines. The history of 100 workers whose loss of employment is attributable to technological advances is reviewed. For each of these individuals, it is determined if he or she was given an alternative job within the same company, found a job with another company in the same field, found a job in a new field, or has been unemployed for 1 year. In addition, the union status of each worker is recorded. The following table summarizes the results.

	Union	**Nonunion**
Same Company	40	15
New Company (same field)	13	10
New Field	4	11
Unemployed	2	5

(a) If the selected worker found a job with a new company in the same field, what is the probability that the worker is a union member?

(b) If the worker is a union member, what is the probability that the worker has been unemployed for a year?

1.98 Group Project: Give each student a bag of chocolate M&Ms. Divide the students into groups of 5 or 6. Calculate the relative frequency distribution for color of M&Ms for each group.

(a) What is your estimated probability of randomly picking a yellow? a red?

(b) Redo the calculations for the whole classroom. Did the estimates change?

(c) Do you believe there is an equal number of each color in a process batch? Discuss.